AD-A034 057 AIR FORCE AVIONICS LAB WRIGHT-PATTERSON AFB OHIO F/G 12/1
SENSITIVITY OPERATORS AND ASSOCIATED SYSTEM CONCEPTS FOR LINEAR--ETC(U)
JUL 76 J G REID
UNCLASSIFIED AFAL-TR-76-118 NL

1 OF 3
AD-A
034 057

SENSITIVITY OPERATORS AND ASSOCIATED SYSTEM CONCEPTS FOR LINEAR DYNAMIC SYSTEMS

AIR FORCE AVIONICS LABORATORY
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

JULY 1976

SENSITIVITY OPERATORS AND ASSOCIATED SYSTEM CONCEPTS FOR LINEAR DYNAMIC SYSTEMS

TARGET ACQUISITION AND STRIKE BRANCH
RECONNAISSANCE AND WEAPON DELIVERY DIVISION

JULY 1976

TECHNICAL REPORT AFAL-TR-76-116

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

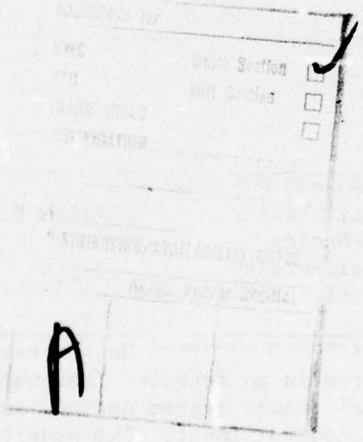
REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFAL-TR-76-118	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SENSITIVITY OPERATORS AND ASSOCIATED SYSTEM CONCEPTS FOR LINEAR DYNAMIC SYSTEMS		5. TYPE OF REPORT & PERIOD COVERED Final Report for Ph.D. Dis. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. Gary Reid Captain, USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFAL/RWT Air Force Avionics Laboratory Wright Patterson AFB, Ohio 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS AFIT0001
11. CONTROLLING OFFICE NAME AND ADDRESS AFIT/ENE Air Force Institute of Technology Wright Patterson AFB, Ohio 45433		12. REPORT DATE July 1976 13. NUMBER OF PAGES 212
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Parameter Sensitivity Linear System Theory Sensitivity Control Parameter Identification Identifiability Input-Design Controllability Eigenvalue Sensitivity Walsh Functions		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Output parameter sensitivity for linear dynamic systems is treated in an operator framework. Parameter sensitivity is derived for a general linear system defined as the sum of linear operators on the initial state and control input. The sensitivity-related properties of insensitivity, sensitivity controllability, parameter identification and identifiability, and optimal control design for both minimum sensitivity and improved identification are then dealt with in terms of the sensitivity operators and their adjoints. This theoretical foundation is applied to		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

give fundamental descriptions of the parameter sensitivities in linear ordinary differential equation systems. The discussion focuses on linear time-invariant systems where a new algebraic representation of the parameter sensitivities is obtained. This representation shows the essential way in which the eigenvalue sensitivities are related to the modes in the sensitivity operators. Its geometrical structure provides algebraic conditions for insensitivity, sensitivity controllability, and both local and structural identifiability. Next the computational aspects are examined by using this matrix-operator form in a quasilinearization algorithm for parameter identification and in optimal control design for both minimum sensitivity and improved parameter identification. Computational simplification are further enhanced in the input design problem by selecting Walsh basis functions for the Rayleigh-Ritz-Galerkin procedure. Throughout the presentation the operator approach is used as a unifying foundation, and the computational results for linear time-invariant systems should be a practical alternative to currently employed sensitivity system differential equation methods.



FOREWORD

I wish to express my sincere appreciation to all those who have assisted me through the doctoral program, the conduct of this research, and the preparation of this report. The list is far too long to mention all, but a special word of thanks goes to Lieutenant Peter Maybeck, my dissertation committee chairman, for his continued personal encouragement and for his invaluable suggestions during the research and preparation of the report. Likewise a special thanks to both Captain Robert Asher and Major James Dillow for the paramount role which they played in my educational training and for their great assistance and careful review in carrying out this research.

Also, on the academic faculty, I wish to express appreciation to Major Gerald Anderson for his assistance as temporary advisor during my formal course work and to Dr. Gary Lamont for his suggestions during a special study in which the very early concepts of this research were formulated and for his critique of the final report. Also I am indebted to Captain Randall Gressang for numerous consultations on linear operator theory and to Captain Eric Lindberg for the suggestion of the example in Section III.2.b.

The majority of this research was carried out while I was assigned with the Air-Air Analysis Group, Analysis and Evaluation Branch, Navigation and Weapon Delivery Division, Air Force Avionics Laboratory. I would like to thank retired Colonel William Delaney, Dr. Jesse Ryles, and, again, Captain Robert Asher for permitting me the time to conduct these investigations.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

List of Illustrations

List of Tables

List of Symbols

The following list of symbols presents the primary symbols and notation used throughout the report. It does not include those symbols which are used just one time for the purpose of illustrating some particular point in passing, nor does it include all of the variations of subscripts, superscripts, or functional dependencies. These variations are explained within the report itself.

Symbol	Meaning	Location First Appearing or Defined
$a_j(t)$	Scalar functions in polynomial representation of $e^{At}_{(i)}$	Eq (131)
$a(t)$	2n dimensioned vector function with components $a_j(t)$	Eq (185)
$\overline{a}^*(t)$	rx2nr dimensioned matrix function with elements $a_j(t)$	Eq (219)
$\hat{a}_k(t)$	Scalar functions used in definition of second order parameter sensitivities	Eq (454)
b_i	Real parameter component	Pg 5
b	Real p dimensional parameter vector	Pg 5
b_o	Nominal parameter vector	Pg 5
$\hat{b}$	Estimate of $b \in R^p$	Sections II.7 and V
$\hat{b}^*$	Optimal estimate of $b \in R^p$	Sections II.7 and V
"b"	Subscript denoting total derivative with respect to $b \in R^p$	Pg 21
c_1, c_2	Real constants	Pg 36, Pg 131
d	Initial state vector	Pg 5
$\overline{d}$	Augmented "sensitivity" initial state	Eq (6)
e^t	Exponential raised to power of superscript	Pg 59
e^{At}	Matrix exponential	Pg 59

Symbol	Meaning	Location First Appearing or Defined
e_i	The i^{th} Euclidean basis vector	Pg 57
$f(x,b,t)$	General nonlinear function	Pg 4
$f(t)$	2n dimensioned vector function	Eq (405)
$f_j(t)$	Components of $f(t)$	Pg 174
$g(t;u)$	2nr dimensioned function	Eq (186)
$g(q)$	Polynomial	Pg 65
$g_i(q)$	Polynomial	Pg 65
$h(t)$	Scalar function of t	Eq (301)
i	Subscript or superscript generally denoting i^{th} parameter component	Pg 5
(i)	Subscript denoting partial derivative with respect to b_i	Pg 5
j	Subscript of superscript	
k	Subscript or superscript. Also scalar constant	Eq (300)
k_{ij}	Scalar constants in weighted trace cost functional	Eq (75)
m	Number of outputs	Pg 20
n	State dimension for ordinary differential equation systems	Pg 5
n_k	Multiplicity of eigenvalue q_k in characteristic polynomial of A matrix	Pg 60
p	Parameter dimension	Pg 5
q_k	k^{th} distinct eigenvalue of A matrix	Pg 60
σ_M	Maximal eigenvalue of V_Q	Pg 150
r	Control dimension	Pg 5
s	Dummy parameter of integration	
s_n	Scalar step size for gradient iterations	Pg 34
s_n^*	Optimal step size	Eq (47)

Symbol	Meaning	Location First Appearing or Defined
t	Independent "time" variable	Pg 4
t_o	Initial time	Pg 4
t_f	Final time	Pg 4
u	Control input	Pg 4
u^*	Optimal control	Pg 28, Pg 51
$v^{(i)}$	Output parameter sensitivity with respect to b_i	Pg 21
$v_j^{(i)}$	Sensitivity of j^{th} output with respect to b_i	Eq (195)
$v^{(i,j)}$	Second order output parameter sensitivity with respect to b_i and b_j	Eq (40)
w_j	j^{th} column vector of Walsh matrix W_a	Pg 159
x	System state	Eq (4)
y	System n-dimensional output	Eq (8)
y_j	j^{th} output component	Pg 83
z	Output error function	Eq (32)
"z.i."	Subscript denoting zero input response	Section II
"z.s."	Subscript denoting zero state response	Section II
A	nxn dimensioned plant matrix	Pg 5
$\bar{A}$	n(p+1)xn(p+1) dimensioned "sensitivity system" plant matrix	Eq (7)
$\tilde{A}_i$	2nx2n matrix of i^{th} sensitivity system	Eq (118)
$A_{K_{z.i.}}$	Zero-input self-adjoint operator	Eq (78)
$A_{K_{z.s.}}$	Zero-state self-adjoint operator	Eq (82)
B	nxr control input matrix	Pg 5

Symbol	Meaning	Location First Appearing or Defined
$\bar{B}$	n(p+1)xr "sensitivity system" control matrix	Eq (7)
C	mxn output matrix	Pg 54
$\bar{C}$	m(p+1)xn(p+1) "sensitivity system" output matrix	Eq (94)
D	State space	Pg 20
$E^{(0)}, E^{(i)}, E, \bar{E}$	Matrices used in matrix operator representation of zero-input sensitivities	Eqs (187)(188)(193)(200)
E_j^*, E^*	Matrices used in zero-input identifiability	Eqs (241)(242)
$F^{(0)}(k), F^{(i)}(k), F^{(i)}, \bar{F}$	Matrices used in the matrix-operator "component" representation of the zero-input sensitivities	Eqs (407)(408)(412)(414)
$G^{(0)}, G^{(i)}, G, \bar{G}$	Matrices used in the matrix-operator representation of the zero-state sensitivities	Eqs (189)(190)(194)(200)
G_j^*, G^*	Matrices used in zero-state identifiability	Eqs (252)(253)
$\bar{G}_c$	A γx2nr dimensioned matrix of rank γ	Pg 91
$H^{(0)}(k), H^{(i)}(k), H^{(i)}, \bar{H}$	Matrices used in the matrix-operator "component" representation of the zero-state sensitivities	Eqs (409)(410)(413)(414)
$[\bar{E}_1 \quad \bar{G}_1]$	A γ_1x2n(r+1) dimensioned matrix of rank γ_1	Pg 92
[F G]	A npx2n(r+1) dimensioned matrix	Eq (278)
J_e, J_{LM}, J_Q	Cost functionals used in parameter identification	Eqs (29)(35)(51)
$J_K, J_{K_{z.i.}}, J_{K_{z.s.}}$	Cost functionals used in input design	Eqs (75)(76)(80)
J_{S_H}, J_S	Cost functionals used in minimum sensitivity control design	Eqs (22)(295)

Symbol	Meaning	Location First Appearing or Defined
K, K_1, K_2	Transformation matrices	Section IV.3
K_n	Weighting matrix	Pg 35
$L_2(t_o, t_f; R)$	Space of square integrable functions on interval $[t_o, t_f]$	Pg 20
$L_c(\cdot,\cdot)$	Space of linear continuous mappings	Pg 20
M	Identifiability matrix	Eq (50)
M_Q	Local information matrix	Pg 45
M_c	A 2nx2n dimensioned Gram matrix	Eq (308)
M_c^*	A 2nrx2nr dimensioned partitioned matrix	Eq (218)
O(t)	Control weighting	Pg 122
P(t)	Self-adjoint operator mapping on control space	Eq (302)
Q	Identification weighting matrix	Pg 36
$\overline{Q}$	Partitioned diagonal weighting matrix	Eq (381)
Q_H	2nx2n matrix	Eq (370)
R	Space of real variables	Pg 5
$\overline{R}(t)$	Riccati matrix	Pg 129
S_H	Hilbert space linear system	Pg 20
S_{LC}	Linear time-invariant ordinary differential equation system	Pg 59
S_{TV}	Linear time-varying ordinary differential equation system	Pg 54
T	Zero-input system operator	Pg 8
T*, etc	Adjoint operator of T	Pg 22
U	Space of controls	Pg 20
V	Total output sensitivity operator	Eq (11)

Symbol	Meaning	Location First Appearing or Defined
V_Q	"Rayleigh-Ritz-Galerkin" matrix	Eq (374)
$\tilde{V}_{i_{k,j}}$	"Component" of $e^{At}_{(i)}$	Eq (128)
W	Zero-state system operator	Pg 20
W_a	Walsh matrix	Eq (392)
X	Augmented "sensitivity system" state	Eq (6)
X_c	Reduced dimension "sensitivity state"	Pg 137
Y	Augmented "sensitivity system" output	Eq (92)
$Z_{k,j}$	"Component" of A	Pg 60
$\tilde{Z}_{i_{k,j}}$	"Component" of $\tilde{A}_i$	Pg 63
α, α^*	Scalar constants	Pgs 152-153
$\alpha_j(t)$	Scalar functions in "minimal" polynomial representation of e^{At}	Eq (115)
β	Number of eigenvalues with non-zero sensitivity for at least one parameter component	Pg 72
β_i	Number of eigenvalues with non-zero sensitivity with respect to b_i	Pg 70
β_M	Maximal eigenvector of V_Q	Pg 150
γ	Rank of $\bar{G}$	Pg 91
γ_1	Rank of $[\bar{E} \quad \bar{G}]$	Pg 92
$\delta(t)$	Unit sampling function	Pg 37
δ_{ij}	Dirac delta function	Pg 67
ε_1	Error term of second order in Δb	Pg 23
ε_2	Error term of third order in Δb	Pg 176
η	Measurement noise process	Pg 31
θ_k	Constant vector	Eq (393)

Symbol	Meaning	Location First Appearing or Defined
κ, ℓ	Indices	Section VII
$\lambda(t)$	Lagrange multiplier vector function	Pg 152
$\mu, \tilde{\mu}_i, \bar{\mu}$	Order of minimal polynomials of $\tilde{A}$, A_i, and $\bar{A}$ respectively	Section III.2
$\mu_k, \tilde{\mu}_{i_k}, \bar{\mu}_k$	Eigenvalue multiplicity of q_k in minimal polynomials of $A, \tilde{A}_i, \bar{A}$	Section III.2
ν	Rank of [F H]	Pg 116
$\xi^{(i)}$	State sensitivity with respect to b_i	Pg 4
π	Projection operator	Eq (307)
ρ	Number of distinct eigenvalues of A	Pg 60
σ_k	Real part of complex eigenvalue, q_k	Pg 171
τ	Number of distinct real eigenvalues	Pg 171
$\phi(t;u)$	2n-dimensioned vector function of t	Eq (371)
$\psi(q), \tilde{\psi}_i(q), \bar{\psi}(q)$	Minimal polynomials of $A, \tilde{A}_i, \bar{A}$	Eqs (112)(126)(158)
ω_k	Imaginary part of complex eigenvalue, q_k	Pg 171
$\Delta(q)$	Characteristic polynomial of A	Eq (111)
Λ	Vandermonde matrix	Pg 171
$\Phi(t,s)$	State transition matrix of A(t)	Pg 56
$\Phi(t)$	State transition matrix of A	Pg 59
$\Psi(t)$	Matrix function	Eq (395)
H	Hamiltonian matrix	Eq (378)
M	Identifiable subspace	Pg 40
$N(\cdot)$	Null space	Pg 23
$R(\cdot)$	Range space	Pg 23
U	Control function space	Pg 23
Y	Output function space	Pg 22

Section I

INTRODUCTION

A central topic of modern system theory is parameter sensitivity. The first step in the analysis of a system is to obtain a mathematical model of that system. Such a mathematical model invariably contains parameters (fixed constants such as the value of a resistor in an electrical network) which are known with varying degrees of certainty. Parameter sensitivity, then, is the study of any property of this mathematical model which might be altered by a change of these parameter values from their nominal or assumed values.

To limit the scope of the discussion, the parameters are assumed to be constant and only local variations of the parameters from their nominal values will be treated. Also, the main concern will be with variations in the mathematical model output (or system response) caused by these local perturbations of the parameters. To be specific, if it is not stated otherwise, then "parameter sensitivity" will be defined as the Fréchet derivative of the mathematical model output with respect to the unknown constant parameters evaluated at the nominal parameter values. This has become a fairly standard definition of parameter sensitivity. (See, e.g., Tomovic and Vukobratovic (Ref 29)).

Notice that a derivative (or partial derivative) is a linear operator (e.g., (Ref 3: 182)) and hence the parameter sensitivities will also be referred to as "sensitivity operators". Indeed, this is the central theme of this presentation: the parameter sensitivities are linear operators and by treating them as such many of their system properties become more apparent and easier to deal with. By further limiting the discussion to linear mathematical models of systems, the system output is defined by

the sum of linear operators on the initial conditions and the control input. The "sensitivity operators" are then defined in terms of the Fréchet derivatives of these system operators. This operator structure is used as a unifying foundation for the sensitivity-related system properties which are presented.

Often the first step in the analysis of a system is to identify or estimate the system parameters from experimental measurements, and parameter sensitivity is the basis of many numerical techniques for parameter identification. The sensitivity operators are intimately related to the question of "identifiability"; that is, the a priori capability to determine whether or not a set of parameters may be uniquely estimated from a given set of experimental measurements. Also, parameter identification capability may be optimized by proper design of the sensitivity operators through either selection of the initial conditions or control input. The sensitivity operators are analogous to signal strength in a communication channel, and it will generally improve estimation capability if the sensitivities are "large".

Once a system model with fairly accurate nominal parameter values is obtained, usually the next step in system analysis is to design a control function to achieve some objective or optimize some performance criterion. Here one generally wishes the system to be forgiving of one's inaccurate knowledge of the true parameter values, and so, in this case, one would like the sensitivity operators to be small. Associated with this objective we discuss the sensitivity-related system properties of insensitivity, sensitivity operator controllability, and optimal design of a control input to minimize the value of the sensitivity operators.

The sensitivity-related system properties mentioned above are considered in this dissertation, but they are examples of just a few of the

many important ways which parameter sensitivity enters into system theory. Each of these properties is ideally treated via an operator viewpoint and many useful digital computational techniques are thereby obtained. Finally, we note that an operator structure is equally applicable in either the time domain or frequency domain, but to limit the scope of the presentation we will only treat the time domain description of linear dynamic systems.

1. BACKGROUND LITERATURE

The earliest interests in parameter sensitivity stemmed mainly from the effect which parameter variations would have upon control system performance. Bode (Ref 4) was one of the first investigators to consider parameter sensitivity in dynamic system analysis. Broadly, he defined the sensitivity function as the ratio between a given percentage change in a parameter component and the resulting percentage change in the system output (delivered voltage in an electrical network). Through frequency domain analysis he showed that feedback could significantly reduce the "sensitivity function" in an electrical network. This original work spurred a great deal of research into a frequency domain analysis of parameter sensitivity and an analysis of the effect of feedback on reducing sensitivity in automatic control systems. (See, e.g., Horowitz (Ref 14) and Horowitz and Shaked (Ref 15).

Although the operator approach to parameter sensitivity is closely related to a frequency domain treatment, this specific analysis treats linear dynamic systems in a time domain state-variable description (Kalman (Ref 16)). Tomovic (Ref 28) develops a time-domain description of parameter sensitivity for linear and nonlinear dynamic systems and introduces the concept of the "sensitivity system"; that is, if a system is described by the nonlinear differential equation

$$\dot{x}(t;\, b) = f(x,\, b,\, t) \qquad x(t_o) = d(b) \qquad t \in [t_o,\, t_f] \tag{1}$$

where b is an unknown constant parameter with nominal value b_o, then the linearized differential equation

$$\frac{d}{dt}\, \xi\,(t;\, b_o) = \frac{\partial f}{\partial x}(x,b,t) \Big|_{b = b_o} \xi(t;\, b_o) + \frac{\partial f}{\partial b}\,(x\,,\, b,\, t) \Big|_{b = b_o} \tag{2}$$

is referred to as the "sensitivity system", and for $t \geq t_o$,

$$\xi(t;\, b_o) \equiv \frac{\partial x}{\partial b}\,(t;\, b) \Big|_{b = b_o} \tag{3}$$

is termed the "sensitivity function" or "parameter influence function".

This fundamental differential equation description has been the basis for nearly all of the time-domain analyses of the parameter sensitivities. (See,e.g., (Ref 29) (Ref 6) as general references on sensitivity). For general nonlinear systems this is perhaps the most appropriate way of dealing with parameter sensitivity, and very few general results beyond this basic relationship have been obtained. However, for linear systems, and in particular for linear time-invariant ordinary differential equation systems, there has been a tremendous amount of research into the structure and system properties of the "sensitivity system". Our specific applications focus on such linear time-invariant ordinary differential equation systems, and so we discuss the relevant literature of parameter sensitivity for such systems. However, before doing so it is important to comment that the fundamental difference between the research reported herein and that of previous investigators is the fact that we do not use a "sensitivity system" differential equation representation of the parameter sensitivities; rather we use an operator time-domain description of the linear system. To obtain the parameter sensitivities we then directly

differentiate this operator description. Therefore, although the previous literature on sensitivity is quite pertinent in that many of the results obtained are similar or identical to those reported here, the fundamental approach taken in deriving these results is generally quite different.

a. Structural Properties

Consider the system

$$\dot{x}(t;u) = A(b)x(t;u) + B(b)u(t) \qquad x(0) = d(b) \qquad t\epsilon[0,t_f] \qquad (4)$$

where x(t) is an n-dimensional state vector, u(t) an r-dimensional control input, and b a p-dimensional real parameter vector which parameterizes the matrices A and B and the initial condition vector d. It is assumed that these latter quantities are continuously differentiable with respect to the individual parameter components b_i at a nominal value of the parameter vector $b_o \in R^p$. Then the state sensitivities ($\xi^{(i)}(t;u) \equiv \frac{\partial x(t;u)}{\partial b_i}$) may be computed from the so-called "sensitivity system" (Ref 29)

$$\dot{X}(t;u) = \bar{A}X(t;u) + \bar{B}u(t) \qquad X(0) = \bar{d} \qquad (5)$$

where all the quantities are evaluated at $b_o \in R^p$ and are defined by

$$X(t;u) \equiv \begin{bmatrix} x(t;u) \\ \xi^{(1)}(t;u) \\ \vdots \\ \xi^{(p)}(t;u) \end{bmatrix}_{n(p+1)x1} \qquad \bar{d} \equiv \begin{bmatrix} d \\ d_{(1)} \\ \vdots \\ d_{(p)} \end{bmatrix}_{n(p+1)x1} \qquad (6)$$

$$\bar{A} \equiv \begin{bmatrix} A & 0 & \cdots & 0 \\ A_{(1)} & A & & \\ A_{(2)} & 0 & A & \\ \vdots & \vdots & & \vdots \\ A_{(p)} & 0 & \cdots & A \end{bmatrix} \qquad \bar{B} \equiv \begin{bmatrix} B \\ B_{(1)} \\ \\ \vdots \\ B_{(p)} \end{bmatrix} \tag{7}$$

The subscript within parenthesis indicates partial differentiation with respect to the parameter component. Therefore, by a direct analysis the complete solution for all of the parameter sensitvities would require solving $n(p + 1)$ coupled linear differential equations; if n and p are both large this can be quite a large number of differential equations indeed.

However, the structure of this "sensitivity system" is quite unique and many interesting facts have been discovered about its properties. A number of researchers have investigated canonical form representations and interrelationships which significantly reduce the required number of differential equations for producing the sensitivity system. (Refs 34-42) Much of this work stems from the "sensitivity-points" techniques introduced by Kokotovic (Ref 37). Varshney (Ref 40) provides an excellent summary of these techniques, and shows that the required number of differential equations for generating the first order sensitivity system may be reduced from $n(p + 1)$ to $n(r + 3)$ where r is the control dimension. He also considers extensions of these techniques to obtain minimal order models of the second order sensitivity operators and the first order sensitivities for linear time-invariant time-delay systems.

More recently the controllability and structural properties of the sensitivity system have been an area of considerable interest. Holtzman and Horing (Ref 69) examine optimal control methods for making the sensitivities zero at the terminal time. This work has initiated an interest in sufficient conditions for uncontrollability of the sensitivty system (Refs 48-53). Through a Jordan canonical form analysis of the sensitivity system plant matrix $\bar{A}$, Guardabassi, et al, (Ref 51) (Ref 52) investigate the structural properties and the controllability of the sensitivity system. In reference 51 they examine the minimal polynomial of $\bar{A}$ and obtain a general formula for eigenvalue sensitivity based upon the Jordan canonical form of $\bar{A}$. They conclude that the eigenvalue multiplicities of $\bar{A}$ are no more than twice that of the original A matrix. They also examine uncontrollability of the sensitivity system and obtain a composite of sufficiency conditions for uncontrollability. In reference 52 they give further consideration to the Jordan canonical form of $\bar{A}$ for the special case in which A has no repeated eigenvalues. They show that the structure of the Jordan canonical form of $\bar{A}$ is intimately related to the eigenvalue sensitivities and that sufficiency conditions for uncontrollability are also related to the eigenvalue sensitivities. A general statement of their results is that the sensitivity system is always uncontrollable if the number of parameters is greater than the number of control inputs. Finally, for the special case of single input systems, they obtain necessary and sufficient conditions for controllability of the sensitivity system. These conditions are independent of the nominal parameter value, $b_o \in R^P$, and, again they are dependent upon the eigenvalue sensitivities.

Gupta and Mehra (Ref 84) and Gupta and Hall (Ref 108) consider controllability of the sensitivity system by a direct analysis of con-

trollability of the matrix pair $(\bar{A}, \bar{B})$. By analyzing linear dependence of vectors in the controllability matrix of $(\bar{A}, \bar{B})$ they obtain sufficiency conditions for uncontrollability. They use the linear dependency of column vectors in the controllability matrix of $(\bar{A}, \bar{B})$ to provide a constructive means to reduce the number of differential equations to the minimum number, and remove some of the undesirable computational features associated with reduction of the sensitivity system via canonical form analysis (Ref 40)

An area which has also used the structural relationships of the sensitivity system for analysis is the question of "insensitivity". (Refs 54-67) Various definitions of insensitivity have been utilized, and Guardabassi et al, (Ref 61) provide an excellent summary of the various definitions and conditions for insensitivity. Also they show the relationships between parametric insensitivity and the related concepts of signal invariance (e.g., (Ref 57)), signal insensitivity (e.g., (Ref 64)), and parametric invariance (Ref 54)). All of these concepts are concerned with system conditions which ensure that parametric or signal perturbations cause no change in the system output. Insensitivity generally connotes zero effect for local or small changes of the parameters, whereas the property of invariance is generally concerned with the stronger condition of zero output perturbation for larger changes of the parameters. We will later be concerned with the weaker condition of parametric insensitivity, but some recent work of Guardabassi, et al (Ref 62) shows that insensitivity of a suitably defined associative system can often imply invariance of the original system. The conditions for insensitivity are well-known (e.g., (Ref 61)), and recent research has concentrated on the geometric construction of feedback control laws which achieve insensitivity (e.g., (Ref 55) (Ref 58) (Ref 59) (Ref 65)). We do not delve into this

latter synthesis area; rather, we only wish to demonstrate the ease by which the operator description of the parameter sensitivities enables one to obtain conditions for insensitivity. Also, we will discuss some of the relationships between insensitivity and identifiability. The connection between non-identifiability and insensitivity has been previously noted (e.g., Bonivento (Ref 55), but no specific results on identifiability which utilize this association have been reported.

b. Optimal Control Laws to Minimize Parameter Sensitivity

The determination of control laws which maintain adequate performance with system parameter changes is a basic problem of control theory and sensitivity analysis. The concepts of sensitivity controllability and insensitivity are related to this problem, and, as mentioned in Section I.1.a, there is current research in the synthesis of control laws which achieve insensitivity. Another popular synthesis technique is the design of minimum sensitivity control laws through the use of optimal control theory. (e.g., (Refs 68 - 78)) The basic concept of such techniques is to place some measure of the sensitivity variables within the optimization cost functional, and then use the methodology of optimal control theory and the Maximum Principle (Pontryagin, et al (Ref 22)) to determine the optimal control law or feedback gain matrices which minimize this cost functional.

For the linear time-invariant system (1) with a quadratic cost functional on the state and state sensitivities, the optimal open-loop control law may be uniquely determined via a Riccati equation (e.g., Kahne (Ref 70)). However, unless minimum order sensitivity models are utilized, the Riccati equation has dimension $n(p + 1) \times n(p + 1)$. The computation time by this method can then be quite large. We investigate this specific open-loop minimum sensitivity problem and show that the operator approach provides

an interesting alternative to the nonlinear matrix Riccati differential equation method of solution.

Closed-loop or feedback control laws are known often to have smaller sensitivity than the nominally equivalent open-loop control law (e.g., Horowitz (Ref 14), Cruz and Perkins (Ref 7)). Therefore, there has also been a considerable amount of research into the design of feedback control laws which minimize parameter sensitivity by using both classical frequency domain and optimal control techniques. For the linear time-invariant system (4) the Riccati equation solution does not produce an optimal closed-loop minimum sensitivity control law because the optimal control is based upon the nominal values of the parameter sensitivities, which are computed under the assumption that the control is open-loop. However, there has been quite a number of attempts to design suboptimal minimum sensitivity feedback control laws based upon this Riccati equations structure. (See, e.g., Lamont and Kahne (Ref 72)). Also, there is considerable interest in the design of minimum sensitivity control laws when there is noise entering into the system as well as parameter variations (e.g., (Ref 77)) or when the control law is adaptively adjusted by using on-line estimates of the parameter values (e.g., Ref 74)). However, to limit the scope of our presentation we only consider the open-loop problem and do not discuss the closed-loop adaptive control problems.

c. Parameter Identification and Sensitivity

The problem of estimating the parameters in a dynamic system can be formulated as an optimization problem in which the parameter values are selected to either minimize or maximize some criterion. Most often this criterion involves the difference between the measured system output and the model system output, where the model output is dependent upon the

estimated parameter values. Generally such a criterion is a nonlinear functional of the parameter values, and an iterative technique must be used to obtain the "optimal" estimates. These iterative techniques inherently involve the computation of the parameter sensitivities; the output parameter sensitivities indicate approximately how much the system output will be changed by small changes of the parameter values.

Many different iterative techniques exist for estimating the parameters, and several excellent survey papers have been written of the use of various optimization criterion (See, e.g., Aström and Eykhoff (Ref 80), Cuenod and Sage (Ref 82) and Eykhoff (Ref 83)). We will be concerned with minimizing a quadratic functional of the output error (the difference between the measured system output and the mathematical model output). Such a functional is compatible with either least squares or maximum likelihood identification. (See, e.g., Gupta and Mehra (Ref 84)). Quasilinearization, also known as modified Newton-Raphson or the method of parameter influence coefficients, is a popular method for determining the minimizing estimates. (Ref 79) (Ref 86) It is less complex than the full second order Newton-Raphson method, but it has nearly equal convergence properties. (See Banks and Groome (Ref 81)). However, due to singularity of the information matrix, a conventional gradient technique must often be used as a starting procedure to the quasilinearization method. The Newton-Raphson, quasilinearization, and gradient methods are all specific examples of general iterative techniques known as stochastic approximation (see, e.g., Saridis (Ref 87) for a survey of stochastic approximation) and all of these methods explicitly involve the parameter sensitivities in their computation.

d. Identifiability

One of the important questions of parameter identification is whether or not the parameters in a system model can be uniquely determined from a given set of experimental observations. This question is the essence of what is called "identifiability" (Lee (Ref 18)). If unique estimates of the parameters cannot be obtained (the system is not identifiable), then either the mathematical model or the experiment itself must be modified. Because it is such a basic question and because the research on parameter identification is so extensive, the literature on identifiability is equally diversified.

Tse and Anton (Ref 103), Aoki and Yue (Ref 88), Staley and Yue (Ref 101), Tse (Ref 102) and other have used a stochastic definition of identifiability in which identifiability is defined by whether or not there exists an estimator which gives convergence in probability (or in mean square) to the true parameter value. The concept of stochastic identifiability is then stronger than the deterministic view of identifiability discussed in the above paragraph (uniqueness of solution), and deterministic identifiability is necessary but not sufficient for stochastic identifiability when there is system measurement and/or system plant noise (Tse and Anton (Ref 103)). For linear time-invariant systems with Gaussian white measurement noise, the conditions of stochastic identifiability lead to the conclusion that the local information matrix must be positive definite to ensure local identifiability (Tse (Ref 102)).

On the other hand, many researchers (e.g., Lee (Ref 18)), Bellman and Åström (Ref 89), Glover (Ref 91) (Ref 92), Glover and Willems (Ref 93), and Martenson (Ref 96)) have adopted a deterministic approach to identifiability similar to that discussed in the first paragraph above and this is, in fact, the approach which we will take. In addition to conditions equivalent to the information matrix being positive definite, these

deterministic approaches have provided considerable structural information concerning identifiability in linear time-invariant ordinary differential equation systems. Bellman and Åström (Ref 89), Mehra (Ref 98), Mayne (Ref 97), Denham (Ref 90), and others examine conditions under which certain canonical forms will be identifiable regardless of the nominal parameter values. Mehra (Ref 98) gives an excellent summary of these conditions.

However, if the parameters of the mathematical model are physical variables it may be inconvenient to utilize such canonical form representations for identification. Additionally, Glover and Willems (Ref 93) point out some inconsistencies which can arise in using such canonical forms. Therefore, using arbitrary parameterizations they consider conditions for both local and global identifiability from the system frequency domain transfer function (zero initial conditions). Then in reference 92 Glover extends the work of Popov (Ref 100) to consider general structural conditions for minimal and nonminal parameterization of linear time-invariant systems. He relates these conditions to structural identifiability in such systems, and shows that the parameter identification problem may become ill-posed if the parameterization is nonminimal.

Martenson (Ref 96) also uses a deterministic approach to identifiability. By extending some concepts of Silverman and Meadows (Ref 25) concerning sufficient conditions for observability in linear time-varying systems, he obtains general algebraic sufficiency conditions for local identifiability in both linear and nonlinear dynamic systems.

Finally, in the next section we will discuss the fact that the control input strongly affects the quality of parameter identification capability, and, indeed, the control input may affect whether or not

the system parameters are even identifiable. Åström and Eykhoff (Ref 80) term such conditions on the control input as conditions for "persistent excitation". Through frequency domain analysis (steady-state conditions) Mehra (Ref 98), Payne and Goodwin (Ref 99), Hoberock and Stewart (Ref 94) and others have examined this question in considerable detail and have determined the number of frequency components which are necessary to ensure identifiability in steady state operation. This interaction between identification and the control input is extremely important. Later we examine identifiability of the zero-state response in linear time-invariant systems, and general structural conditions for identifiability are obtained. However, these structural conditions are only necessary, as sufficiency is dependent upon control input conditions for "persistent excitation".

e. Sensitivity Design for Optimal Parameter Identification

The relationship between the control input function and parameter estimation capability has long been recognized. A considerable number of papers have appeared on various aspects of designing control inputs to optimize or improve parameter estimation capability, and these techniques have been successfully applied to many practical problems. (e.g., (Refs 104 - 122)) Mehra (Ref 114) gives an excellent summary of the literature on input design, discussing methods of solution in both the frequency domain and the time domain. Rather than repeat this general discussion, we will discuss only the work which is directly relevant to our approach.

Our approach is entirely motivated by the novel work of Mehra (Ref 112). He shows that the optimal input function to maximize a weighted trace of the information matrix for linear time-invariant ordinary differential equation systems is the eigenfunction corresponding to the largest eigenvalue of a positive self-adjoint operator. He suggests various methods

of solution including the Rayleigh-Ritz-Galerkin method. This is the solution method we later adopt for our operator approach, but Mehra concentrates on a method which transforms the problem into a two-point boundary value problem using the Pontryagin Maximum Principle (Ref 22). Using the sensitivity system differential equations he suggests solving this problem by means of a matrix Riccati differential equation technique. In (Ref 120) a computer algorithm using this Riccati equation approach is outlined, and the solution method is applied to the determination of optimal input functions for the identification of aircraft stability and control derivatives.

This solution method is then further refined by Gupta and Hall (Ref 108) who eliminate much of the undesirable features of numerical integration of the nonlinear matrix Riccati differential equation. They also incorporate the controllability properties of the sensitivity system to reduce the amount of computations still further. The computational aspects of this solution method will be discussed further in Section VII and compared with the new solution technique which is developed using the operator description of the parameter sensitivities and the Rayleigh-Ritz-Galerkin method.

We comment that a recognized shortcoming of the original approach of Mehra (Ref 112) is the use of the weighted trace of the information matrix as the optimization criterion. Unless the weighting matrix is selected judiciously, this optimization criterion can lead to erroneous results and even to conditions in which the local information will be singular (the system will be nonidentifiable). (See,e.g., Reid (Ref 119), Mehra (Ref 114), Zarrop and Goodwin (Ref 122)). Others have suggested that either the trace of the inverse information matrix (e.g., Reid (Ref 119), Gupta and Hall (Ref 108), and Goodwin, et al (Ref 107)) or the determinant of the inverse

information matrix (e.g., Mehra (Ref 114) and Nahi and Napjus (Ref 117)) is a better criterion for optimization. Through the Cramer-Rao lower bound the former criterion is directly related to the variance of the minimum estimation error whereas the latter is directly proportional to the volume of the error ellipsoid of parameter estimation error. (See Nahi and Napjus (Ref 117) for further detailed discussion of various optimization criteria useful in the input design problem).

However, through some results recently obtained by Mehra (Ref 113), both Mehra (Ref 114) and Gupta and Hall (Ref 108) derive iterative algorithms to successively adjust the weighting matrix in the information matrix criterion so that the resulting optimal input minimizes either the trace or the determinant of the inverse of the information matrix. The original results of Mehra (Ref 112) are, therefore, an integral part of these iterative algorithms. Indeed, the maximization of the weighted trace of the information matrix is the most time consuming part of these algorithms. Therefore, although our approach is closely related to reference 112 in using the weighted trace of the information matrix, it has application to other optimization criteria through these more recent results.

2. OBJECTIVES AND ORGANIZATIONS

There are two major objectives in this research and the presentation of this report. The first is to demonstrate a general operator framework for treating parameter sensitivity in linear dynamic systems and to show the usefulness of this operator formulation for analyzing various sensitivity related system properties. The second objective is to utilize this operator formulation to obtain algorithms and computational techniques which will be useful to systems engineers dealing in modeling, identification and control. In particular, our specific computational results mainly emphasize

large scale, linear, multivariable, time-invariant ordinary differential equation systems.

To help achieve the first objective, concepts are initially developed in a very general operator framework, and then these basic concepts are illustrated by specific application to linear ordinary differential equation systems with emphasis on the time-invariant system. Therefore, Section II defines parameter sensitivities for a linear system described in terms of the sum of bounded linear operators on the Hilbert space of initial conditions and the Hilbert space of control inputs. Once defining these sensitivities in terms of the Fréchet partial derivatives of the system operators, we discuss, in turn, the sensitivity-related system properties on insensitivity, sensitivity controllability, open-loop minimum sensitivity control design, parameter identification, identifiability, and sensitivity design to optimize parameter estimation capability. These system properties are chosen for discussion not only because of their system importance, but also because of their direct dependence upon the operator description of the parameter sensitivities.

Section II establishes the fundamental operator methodology and many of the basic definitions which are subsequently used to consider parameter sensitivity in linear ordinary differential equation systems (Section III - VII). This general operator presentation allows an appealing compactness of notation in the development concepts, and it is written under the assumption that the reader has a basic familiarity with normed linear spaces and some of the elementary concepts of functional analysis and linear operator theory. Luenberger (Ref 20) is an excellent background reference for nearly all of the material which is discussed. However, the reader who lacks this mathematical background should not be discouraged from reviewing Section II, as this section contains much of the physical

and geometric motivation for the properties with which we deal. Rather, such a reader is encouraged to skim the mathematics lightly while paying closer attention to the general discussion and definitions.

Section III develops a general operator description of the parameter sensitivities for linear ordinary differential equation systems. In such systems the key to making the operator formulation a useable approach lies in taking the partial derivative of the state transition matrix with respect to the parameter components. General results are obtained for time-varying plant matrices. Then for time-invariant plant matrices a new algebraic description of these partial derivatives is derived. This description gives considerable structural insight and produces practical computational alternatives to conventional "sensitivity system" differential equation approaches.

In Section IV the geometric aspects of this representation are pursued by considering the properties of insensitivity, sensitivity controllability, and parameter identifiability. The operator approach of Section III is utilized, and new general algebraic conditions are obtained for local and structural identifiability.

The computational aspects of this algebraic representation are next examined by considering a quasilinearization algorithm for parameter identification (Section V), optimal open-loop control design for minimum sensitivity (Section VI), and sensitivity design for optimal parameter identification (Section VII). In each case, detailed algorithms are presented, and the computational aspects of these methods are compared with conventional procedures using the sensitivity system and the Maximum Principle. To help illustrate the computational techniques some simple examples are presented within the respective sections.

Section II

LINEAR SYSTEMS ON A HILBERT SPACE

In this section parameter sensitivity operators and their system properties are considered for a very general linear system, S_H. This system is defined in terms of bounded linear operators on real Hilbert spaces. For notational convenience and to simplify the discussion, only the system output and output sensitivity operators are discussed; however, a similar operator description of the system state and state sensitivities could also be provided.

The intent of this section is to help motivate the use of an operator representation for the treatment of parameter sensitivity in general linear systems. This discussion provides the foundation for our later analysis of parameter sensitivity in linear ordinary differential equation systems (Sections III - VII). However, besides providing a unifying framework for this later discussion, it is important to note that this general treatment should facilitate the extension of these sensitivity concepts to such linear systems as general hereditary differential systems and parabolic partial differential equation systems. These systems can be put in an operator - Hilbert space setting (see, eg, Delfour and Mitter (Ref 9) (Ref 10) or Baras, et al, (Ref 1)), and such mathematical models are beginning to play an increasingly important role in applications of systems theory to chemical, biological, economical, and other large scale, complex systems.

The organization of this section is the following: In Section II.1 the Hilbert-space system S_H is defined and then in Section II.2 some notational conventions are described. In Sections II.3 and II.4, and

II.5 the control related properties of insensitivity, sensitivity controllability, and optimal design for minimum sensitivity, respectively, are defined and discussed. Finally, in Sections II.6, II.7 and II.8, identification related sensitivity properties are described. In particular, these sections discuss general identification algorithms, identifiability, and optimal sensitivity design for improved identification, respectively.

1. SYSTEM DEFINITION

The real m-dimensional mathematical model output of the system S_H is uniquely defined on the closed interval $[t_o, t_f]$ by the operator expression

$$y(t; b, u) = T(t; b)d(b) + W(t; b)u \quad t\varepsilon[t_o, t_f] \tag{8}$$

where d(b) is the initial state of the system from the Hilbert space D; $u \varepsilon L_2(t_o, t_f; U)$ is the control function input over the interval $[t_o, t_f]$ and U is the Hilbert space of controls; $T(t; b) \varepsilon L_c(D, R^m)$ is the continuous zero-input linear operator; $W(t; b) \varepsilon L_c(L_2(t_o, t; U), R^m)$ is the casual, continuous zero-state linear operator; and $b \varepsilon R^p$ is a p-dimensional real parameter vector which parameterizes the initial state and both the zero-input and zero-state system operators. The components of the parameter vector b are designated b_i, i = 1, 2, ...p, and the nominal a priori value of the $b \varepsilon R^p$ is designated $b_o \varepsilon R^p$. The mathematical model will generally be assumed evaluated at the nominal parameter values; therefore, for convenience we will henceforth delete from the notation the explicit dependency of the output and various system operators upon b_o, and this dependency will only be shown if $b \neq b_o$ or if it is required for some reason of emphasis.

At the nominal $b_o \in R^p$ and for each $t \in [t_o, t_f]$ it is assumed that d, T(t) and W(t) are continuously Frechet differential with respect to each parameter component b_i, i = 1, 2, ...p. These partial derivatives are denoted $d_{(i)}$, $T_{(i)}(t)$, and $W_{(i)}(t)$. Since these partials are continuous, the total Frechet differential is (Ref 3: 195)

$$d_b \Delta b = \sum_{i=1}^{p} d_{(i)} \Delta b_i \tag{9}$$

where $\Delta b = b - b_o$ and similarly for $T_b(t)$ and $W_b(t)$.

The component sensitivity operators are defined as the Frechet partial derivatives of the system output with respect to each parameter component, b_i, $v^{(i)}(t; u) = y_{(i)}(t; u)$. Because of our assumptions these become[1]

$$v^{(i)}(t; u) = (T_{(i)}(t)d + T(t)d_{(i)}) + W_{(i)}(t)u \tag{10}$$

The total sensitivity operator is denoted by the partitioned operator

$$V(t; u) = [v^{(1)}(t; u)\ v^{(2)}(t; u)\ \ldots\ v^{(p)}(t; u)]_{mxp} \tag{11}$$

so that the sensitivity "differential" becomes

$$V(t; u)\,\Delta b = \sum_{i=1}^{p} v^{(i)}(t; u)\,\Delta b_i \tag{12}$$

As a final comment, since we are dealing with a linear system, we are assured that we may uniquely decompose the output of the system into its zero-input, $y_{z.i.}(t) = T(t)d$, and zero-state portion, $y_{z.s.}(t; u) = W(t)u$. (Ref 33) The parameter sensitivities may be

---

[1]Notice that this "chain rule" for the derivative applies because the individual partial derivatives are continuous and because $b_i \in R$. (Ref 3: 187-188). If b were assumed to be from a general normed linear space, then a derivative in this form would not apply.

similarly decomposed, and we will use such a decomposition frequently in the discussion and the investigation of the various sensitivity related system properties.

2. NOTATIONAL CONVENTIONS

We will often be interested in the system output and output sensitivities over the entire time interval $[t_o, t_f]$. The space of output functions over this interval will be denoted $\mathcal{Y} = L_2(t_o, t_f; R^m)$. An element from this function space will be represented by $y(\cdot)$ or y. If the time interval is restricted to $[t_o, t]$, then the function space will be denoted $\mathcal{Y}_t$. Similar script notations will be utilized for the control function spaces, $\mathcal{U} = L_2(t_o, t_f; U)$ and $\mathcal{U}_t = L_2(t_o, t; U)$.

In the discussion of the sensitivity-related properties the adjoint operator is frequently encountered. The adjoint of a linear operator will be denoted by that linear operator with a superscript "*". For example, the adjoint of $T(t) \in L_c(D, R^m)$ is denoted by $T^*(t) \in L_c(R^m, D)$, and is defined by the inner product relation

$$[y(t)/T(t)d] = [T^*(t)y(t)/d] \tag{13}$$

It will be assumed that the reader has a basic familiarity with the adjoint operator and its properties, and Luenberger (Ref 20) is recommended as an excellent background reference. However, the following relationship is used so often that it is stated here for convenience:

LEMMA 1

Let $A \in L_c(H,K)$ where H and K are real Hilbert spaces. Then

$$\text{i)} \quad (R(A))^{\perp} = N(A^*) \tag{14}$$

$$\text{ii)} \quad \overline{R(A)} = (N(A^*))^{\perp} \tag{15}$$

See Luenberger (Ref 20: 157) for the proof of this lemma. Note that $R(\cdot)$ and $N(\cdot)$ denote the range and null spaces of the linear operator, the superscript $\perp$ is the orthogonal complement space, and the overbar here denotes closure of the subspace. Also note that many of our later manipulations with the adjoint operator are simplified considerably because we have assumed the underlying spaces to be real Hilbert spaces, and such spaces are dual to themselves.

We now proceed with the development of the various sensitivity-related system properties. In developing these properties the following first-order Taylor's series equation for the output function $y(b_o + \Delta b)$ plays a fundamental role:

$$y(b_o + \Delta b) = y(b_o) + V\,\Delta b + \varepsilon_1$$

$$= y(b_o) + \sum_{i=1}^{p} v^{(i)}\,\Delta b_i + \varepsilon_1 \qquad (16)$$

where

$$\lim_{||\Delta b|| \to 0} ||\varepsilon_1|| \,/\, ||\Delta b|| = 0$$

It will generally be assumed that design is accomplished at a nominal parameter value, $b_o \in R^p$, but that the true parameter value, $b \in R^p$ (not necessarily equal to b_o), is unknown. The first three system properties discussed are parameter insensitivity, sensitivity controllability, and optimal open-loop control design for minimum sensitivity. Each of these is related to making the sensitivity operator, V, small, and hence all have the general objective of reducing the perturbation in system response

$$\Delta y = y(b_o + \Delta b) - y(b_o) \qquad (17)$$

caused by small changes of the system parameter, $\Delta b = b - b_o$. On the other hand, the last three properties discussed are system parameter identification, identifiability and sensitivity design for improved parameter estimation; here it will generally improve parameter identification capability if the sensitivity operators are large rather than small. The Taylor's series relationship is used in developing each of these sensitivity concepts, and the operator formulation is used as a unifying foundation for the discussion.

3. PARAMETER INSENSITIVITY

Background literature on parameter insensitivity was discussed in Section I.1.a. Few specific results can be made at this level of generality (for the Hilbert space system S_H), and so the objective here is merely to establish the definitions which will be used in Section IV for obtaining algebraic insensitivity conditions in linear time-invariant ordinary differential equation systems. The main difference between these definitions and previous ones is the separation between zero-input and zero-state insensitivity[1]:

Definition 1

The system S_H is b_i -- component zero-input insensitive if $v^{(i)}_{z.i.}(t) = 0$ for all $t \in [t_o, t_f]$. It is totally zero-input insensitive if this is true for all $i = 1, 2, \ldots p$.

Definition 2

The system S_H is b_i -- component zero-state insensitive if $v^{(i)}_{z.s.}(t; u) = 0$ for all $t \in [t_o, t_f]$ and all u. It is totally zero-state insensitive if this is true for all $i = 1, 2, \ldots p$.

---

[1]Note that zero-state insensitivity is what Guardabassi, et al, (Ref 61) refer to as "hypoinsensitivity".

4. SENSITIVITY CONTROLLABILITY

A central topic of modern system theory is controllability (e.g., Kalman, et al, (Ref 17)). Controllability properties of the "sensitivity system" in linear time-invariant ordinary differential equation systems has been an area of considerable research because of the relation to making the sensitivities identically zero at the terminal time and for producing minimal order differential equation models of the sensitivities (see Section I.1.a for a discussion of the relevant background literature). We treat sensitivity controllability for such systems in Section IV.3. Therefore, like the previous section, our main purpose here is to establish the definitions which will later be used to help obtain practical computational results in Section IV.3.

To conform with our operator treatment of the output parameter sensitivities, we define sensitivity controllability in terms of whether or not the zero-state operators map the control function space $\mathcal{U}$ onto the output space R^m. (See, e.g., Weiss (Ref 32) or Delfour and Mitter (Ref 10).) Recall that

$$v^{(i)}(t;\, u) = v^{(i)}_{z.i.}(t) + W_{(i)}(t)\, u \tag{18}$$

where

$$v^{(i)}(t;\, u) \in R^m \quad \text{and} \quad W_{(i)}(t) \in L_c(\mathcal{U}_t,\, R^m).$$

Therefore, if $W_{(i)}(t)$ maps onto R^m, for any given value of $v^{(i)}(t;\, u) \in R^m$ there exists a control $u \in \mathcal{U}$ such that

$$(v^{(i)}(t;\, u) - v^{(i)}_{z.i.}(t)) = W_{(i)}(t)\, u, \tag{19}$$

and so the sensitivity $v^{(i)}$ is termed "controllable" at time $t \in [t_o,\, t_f]$.

Sensitivity controllability has previously been examined for linear time-invariant ordinary differential equation systems, and generally it has been discussed in the context of controllability (or uncontrollability) of the augmented "sensitivity system", equation (5). Therefore, we define:

Definition 3

The system S_H is output sensitivity controllable at time $t \in [t_o, t_f]$ if the "partitioned" operator

$$\overline{W}(t) \equiv \begin{bmatrix} W(t) \\ W_{(1)}(t) \\ \vdots \\ W_{(p)}(t) \end{bmatrix} \tag{20}$$

maps U_t onto the product space $R^{m(p+1)}$.

Definition 4

The system S_H is output sensitivity uncontrollable if there exists no $t \in [t_o, t_f]$ such that $\overline{W}(t)$ maps U_t onto $R^{m(p+1)}$.

Since the output space is finite dimensional, output sensitivity controllability may be determined from the $m(p+1) \times m(p+1)$ symmetric "sensitivity controllability matrix", $\overline{W}(t)\overline{W}^*(t)$, where $\overline{W}^*(t) \in L_c(R^{m(p+1)}, L_2(t_o,t;U))$ is the adjoint of $\overline{W}(t)$. This well-known result is reiterated in the following lemma (Ref 10):

Lemma 2

i) S_H is output sensitivity controllable at $t \in [t_o, t_f]$ if and only if $\overline{W}(t)\overline{W}^*(t)$ is invertible.

ii) The sensitivity controllable subspace of $R^{m(p+1)}$ is

$$R(\overline{W}(t)) = R(\overline{W}(t)\overline{W}^*(t)) \tag{21}$$

5. CONTROL DESIGN FOR MINIMUM TRAJECTORY SENSITIVITY

In the previous two sections, the properties of insensitivity (zero sensitivity) and sensitivity controllability were discussed. If a system is sensitivity controllable, then a control law may be determined such that the parameter sensitivities will all be zero at the terminal time. (Kalman, et al, (Ref 17).) However, the structural conditions for insensitivity or sensitivity controllability often cannot be achieved (see Section IV). An alternate approach for reducing sensitivity is to design an optimal control law which minimizes some measure of the parameter sensitivities while achieving other system objectives. This optimal minimum sensitivity control problem is the topic discussed in this section.

Background literature on the minimum sensitivity control design was discussed in Section I.1.b. These approaches have relied on the augmented sensitivity system in which the parameter sensitivities are treated as added "pseudo-state" of the system. For linear systems with a quadratic cost functional the optimal open-loop minimum sensitivity control may then be determined explicitly via a Riccati equation technique; however, the drawback is that there are large numbers of "states" and this can cause computational problems. (e.g., Kahne (Ref 70).)

In this section, we formulate an operator, gradient approach to minimum sensitivity control design. By using well-known methods of functional analysis, the minimum sensitivity control problem is transformed into a minimum norm control problem on the control space, $L_2(t_o, t_f; U)$. This problem may then be solved by standard gradient iteration methods. Such an approach has the potential of reducing the computational requirements of the open-loop minimum sensitivity control problem. (See Section VI for application to linear ordinary differential equation systems.)

For simplicity, we assume that we wish to minimize the quadratic functional

$$J_{S_H}(u) = [Y(u)/SY(u)] + [u/u]$$

$$= \int_{t_o}^{t_f} [Y(t;u)/S(t)Y(t;u)]dt + \int_{t_o}^{t_f} [u(t)/u(t)]dt \tag{22}$$

where $Y(u) \in L_2(t_o, t_f; R^{m(p+1)})$ is the augmented output vector function

$$Y(u) = \begin{bmatrix} y(u) \\ v^{(1)}(u) \\ \vdots \\ v^{(p)}(u) \end{bmatrix} \tag{23}$$

and the matrix $S(t)$ is assumed to be non-negative definite and symmetric for each $t \in [t_o, t_f]$.

We may write $Y(u) \in L_2(t_o, t_f; R^{m(p+1)})$ as

$$Y(u) = Y_{z.i.} + Y_{z.s.}(u) = Y_{z.i.} + \bar{W}u \tag{24}$$

where $\bar{W}$ is defined by equation (20). Denoting the adjoint operator of $\bar{W}$ as $\bar{W}^*$ then the cost functional $J_{S_H}(u)$ takes the form

$$J_{S_H}(u) = [u/(\bar{W}^*S\bar{W} + I)u] + 2\,[u/\bar{W}^*Y_{z.i.}] + [Y_{z.i.}/Y_{z.i.}] \tag{25}$$

Since S is assumed positive, the self-adjoint operator

$$(\bar{W}^*S\bar{W} + I) \in L_c(U,U) \tag{26}$$

is invertible. Therefore, the unique minimizing control is given by

$$u^* = -(\bar{W}^*S\bar{W} + I)^{-1}\,\bar{W}^*Y_{z.i.} \tag{27}$$

(Ref 23). However, it would generally be difficult to obtain the inverse of the operator $(\bar{W}^*S\bar{W} + I)$ directly[1], and so for the "open-loop" problem some iterative solution technique may be applied. The gradient of $J_{S_H}(u)$ is

$$\nabla J_{S_H}(u) = 2(\bar{W}^*S\bar{W} + I)u + 2\bar{W}^*Y_{z.i.} \tag{28}$$

and so either a gradient or conjugate gradient algorithm might be utilized to iteratively compute the optimal $u^* \epsilon U$. (e.g., Luenberger (Ref 20).)

In Section VI, we apply this operator method of solution to linear ordinary differential equation systems. Computational considerations as well as some variations of this basic problem will there be discussed.

6. SENSITIVITY OPERATORS IN PARAMETER IDENTIFICATION

The remaining system properties which we investigate are system parameter identifiability and "input design" for improved parameter identification. To motivate these topics this section briefly discusses the fundamental role which the sensitivity operators play in parameter identification.

a. Output Error Functional

Consider the system S_H and assume that $b \in R^p$ is an unknown constant parameter vector which we wish to estimate based upon a measured system

---

[1]If the inverse of $(\bar{W}^*S\bar{W} + I)$ were to be obtained (for example, by Riccati equation techniques - see Section VI) then expression (27) for u^* has the form, at least, of a "closed-loop" control law. However, since the augmented vector $Y_{z.i.}$ is a function of the a priori zero-input sensitivities (which are computed under the assumption that the control is "open-loop"), this minimum sensitivity optimal control cannot be made "closed-loop". This is a well-known dilemma (e.g., Price and Deyst (Ref 75)), and is discussed further in Section VI.

output function which we denote by $y \in \mathcal{Y}$. The control input, u, is assumed to be fixed and completely known. Therefore, to emphasize that the control is fixed, in this section and in Section II.7, we will delete the explicit dependency of the system output upon the control, u. We let $\hat{b}_o$ denote an a priori estimate of $b \in R^p$, and $y(\hat{b}_o) \in \mathcal{Y}$ denotes the corresponding mathematical model output function. The system sensitivities, $v^{(i)}$, $i = 1, 2, \ldots p$, are then assumed to be evaluated along the mathematical model output, and they are assumed to exist and be continuous for all $\hat{b} \in R^p$.

The estimation criterion which we adopt is the commonly used (e.g., (Ref 80)) output error functional

$$J_e(\hat{b}) = [y - y(\hat{b})/y - y(\hat{b})] \tag{29}$$

The best estimate, $\hat{b}^*$, is then the $\hat{b}$ which minimizes $J_e(\hat{b})$. Note that other more generalized cost functionals might be selected in order to account for measurement noise or system process noise. This generalization is considered in Section II.6.d, and it does not change the basic approach.

b. Linearized Measurement Equation and Quasilinearization

Since the existence and continuity of the sensitivity operators at $\hat{b}_o$ are assumed, we may linearize the true system output, y, about the nominal output, $y(\hat{b}_o)$, through the generalized Taylor's formula as follows:

$$\begin{aligned} y &= y(\hat{b}_o) + V\Delta b + \varepsilon_1 \\ &= y(\hat{b}_o) + \sum_{i=1}^{p} v^{(i)} \Delta b_i + \varepsilon_1 \end{aligned} \tag{30}$$

where $\Delta b = b - \hat{b}_o$ and

$$\lim_{||\Delta b|| \to 0} ||\epsilon_1|| \,/\, ||\Delta b|| = 0 \tag{31}$$

Defining the output error

$$z = y - y(\hat{b}_o) \tag{32}$$

we obtain the linearized measurement equation

$$z = V\Delta b + \epsilon_1 \tag{33}$$

If there is additive measurement noise, $\eta \in Y$, then the linearized measurement equation becomes

$$z = V\Delta b + \epsilon_1 + \eta \tag{34}$$

This equation, although quite simple, is of fundamental importance to our later development of identifiability criteria and input design; it is this equation which clearly demonstrates the importance of the sensitivity operators in the parameter identification problem. To first order, the sensitivity operators tell one how much the system output will be perturbed by small changes in the system parameters. If the sensitivity is high, then, in some sense, one can more accurately estimate the true parameter values.

Now assume the error term ϵ_1 (which is of second order in Δb) is either small or can be compensated for through iterations. Then the output error functional becomes

$$J_{LM}(\Delta b) = [z - V\Delta b / z - V\Delta b] \tag{35}$$

Since Y and R^p are Hilbert spaces and $V \in L_c(R^p, Y)$, the theory of least squares (see Luenberger (Ref 20: 160)) may be used to give the unique minimizing solution

$$\Delta b^* = (V^*V)^{-1} V^* z \tag{36}$$

provided that $(V^*V)^{-1}$ exists. Note that, V^*V is the p x p Gram matrix

with i^{th} row and j^{th} column element

$$V^*V(i,j) = [v^{(i)}/v^{(j)}] \tag{37}$$

while V^*z is the p x 1 vector with i^{th} element

$$V^*z\ (i) = [v^{(i)}/z] \tag{38}$$

Thus it is relatively easy to compute the minimizing solution Δb^*. Also since V is a linear continuous mapping from the finite dimensional space, R^p, its image space must be finite dimensional in Y, and hence must be closed. Then a necessary and sufficient condition for $(V^*V)^{-1}$ to exist is that V be a one-to-one mapping (Ref 10). These statements form the basis of our identifiability discussion of the next section.

The updated parameter estimate is then determined from the equation

$$\begin{aligned}\hat{b}_1 &= \hat{b}_o + \Delta b^* \\ &= \hat{b}_o + (V^*V)^{-1} V^* z\end{aligned} \tag{39}$$

The system output may be linearized about $\hat{b}_1$ and $y(\hat{b}_1)$, and the process of obtaining a new Δb^* repeated. Such a procedure is a generalization of the method of quasilinearization and has found wide-spread application in system parameter identification (see Section I.1.c).

The linearized measurement equation (34) and the corresponding least squares solution, equation (36) provide the greatest motivation for our subsequent discussion of identifiability and input design; however, it is instructive to note the role which the sensitivity operators play in other iterative methods for minimizing the output error functional, $J_e(\hat{b}^*)$, equation (29). In the next section we briefly discuss first and second order gradient procedures, and compare these algorithms with the

quasilinearization algorithm. In all three cases, the sensitivity operators have fundamental importance in the iterations.

c. Gradient Minimization Procedures

Once again consider the general output error functional, $J_e(\hat{b})$, equation (29), and assume that both the first order sensitivities $v^{(i)}$, i = 1, 2, ...p, and the second order sensitivities

$$v^{(i,j)} = \left.\frac{\partial^2 y(b)}{\partial b_i \partial b_j}\right|_{b = \hat{b}_o} \qquad i, j = 1, 2, \ldots p \tag{40}$$

exist and are continuous for all $\hat{b}_o \in R^p$. Then it is straight-forward to show that the first and second order gradients of $J_e(\hat{b})$ are

$$\nabla J_e(\hat{b}) = -2V^* z \tag{41}$$

$$\nabla^2 J_e(\hat{b}) = 2(V^* V - V_b^* z) \tag{42}$$

where V^*z and V^*V are defined by equations (38) and (37), respectively, and the symmetric p x p matrix $V_b^* z$ has i^{th} row and j^{th} column element

$$V_b^* z(i,j) = [v^{(i,j)}/z] \tag{43}$$

Sufficient conditions for $\hat{b}^* \in R^p$ to be a local minimizing solution of $J_e(\hat{b})$ are well-known and are stated in the following lemma:

Lemma 3

For $\hat{b}^* \in R^p$ suppose that $\nabla J_e(\hat{b}^*) = 0$ and that the symmetric p x p matrix $\nabla^2 J_e(\hat{b}^*)$ is positive definite. Then $\hat{b}^*$ is a local minimum of $J_e(\hat{b})$.

Comments

1. Since $y(\hat{b})$ may be a general nonlinear function of $\hat{b}$, no more than a local minimizing solution may be guaranteed; that is, the output

error functional $J_e(\hat{b})$ may have more than one local minimum.

2. If $z = y - y(\hat{b})$ is allowed to take on any value, then it is necessary that the matrix V^*V be positive definite to ensure that $\nabla^2 J_e(\hat{b})$ is positive definite. In Section II.7 we will see that this fact is intimately related to the conditions for local identifiability.

A first order gradient algorithm to find the minimizing $\hat{b}$ may now be defined by the sequence of steps (see, for example, Luenberger (Ref 20: 283 - 287))

$$\hat{b}_{n+1} = \hat{b}_n - \frac{s_n}{2} \nabla J_e(\hat{b}_n)$$

$$= \hat{b}_n + s_n V^* z \tag{44}$$

The gradient, $2V^*z$, is evaluated at $\hat{b}_n$, and the scalar step size s_n is chosen to minimize $J_e(\hat{b}_n + s_n V^* z)$. Through the first order linearization

$$y(\hat{b}_n + s V^* z) \approx y(\hat{b}_n) + s VV^* z \tag{45}$$

we see that

$$J_e(\hat{b}_n + s V^* z) \approx [z/z] - 2s[z/VV^* z] + s^2[VV^* z/VV^* z] \tag{46}$$

Thus, the approximate minimizing s_n is given by

$$s_n^* = \frac{[V^* z/V^* z]}{[V^* z/(V^*V)V^* z]} \tag{47}$$

provided that the denominator is non-zero. For arbitrary $V^*z \neq 0$ (recall that $z = y - y(\hat{b}_n)$) a sufficient condition for the denominator to be non-zero is that the symmetric p x p matrix, V^*V, (evaluated at $\hat{b}_n$) be positive definite. The gradient steps are then repeated until the

gradient vector V^*z, has a norm less than some preselected small positive value. Using the step size s_n^*, the method is known as steepest descent (Luenberger (Ref 20: 286)).

Similarly, the Newton-Raphson algorithm may be used to iteratively compute the minimizing $\hat{b}^* \in R^p$, (see, for example, Luenberger (Ref 20: 284)). Such a procedure is defined by the sequence of steps

$$\hat{b}_{n+1} = \hat{b}_n - (\nabla^2 J_e(\hat{b}_n))^{-1} \nabla J_e(\hat{b}_n)$$

$$= \hat{b}_n + (V^*V - V_b^* z)^{-1} V^* z \tag{48}$$

provided that the indicated inverse exists. Again, for arbitrary $V_b^* z$, a necessary condition for the second order gradient matrix, $\nabla^2 J(\hat{b}_n)$, to be positive definite is that the symmetric p x p matrix V^*V (evaluated at $\hat{b}_n$) be positive definite. Once again the sensitivity operators play a fundamental role in the iterations.

Now compare the first order gradient, the quasilinearization, and the Newton-Raphson algorithms. All three have the same form:

$$\hat{b}_{n+1} = \hat{b}_n + K_n V^* z \tag{49}$$

For the first order gradient algorithm, the weighting matrix K_n is the scalar step size s_n which may be selected according to equation (47). For quasilinearization, K_n is the symmetric p x p matrix, $(V^*V)^{-1}$, and for the Newton-Raphson algorithm it is the second order, symmetric, p x p, gradient matrix $(V^*V - V_b^* z)^{-1}$. As stated previously, we are primarily interested in the linearized measurement equation (34) and the associated quasilinearization algorithm, but it is interesting to note the similarities between the quasilinearization method and these other procedures. In all three cases, the sensitivity operators play a key role in the iterations.

It is also instructive to note the appearance in all three algorithms of the symmetric, p x p, positive matrix, V^*V. Suppose that we use the approximate optimal s_n^* as given by equation (47) in the first order gradient algorithm. Then in each of the three algorithms we require V^*V to be positive definite in order that we may ensure a priori a local solution to each iterative step (considering $z \in Y$ to take on any value). In Section II.7 the matrix V^*V is shown to be of basic importance to the question of local identifiability of the system parameters. Hence, we term this p x p symmetric matrix the "local identifiability matrix" and give it the symbol

$$M(\hat{b}_o) = V^*(\hat{b}_o)V(\hat{b}_o) \tag{50}$$

The dependency upon $\hat{b}_o$ is shown explicitly here for emphasis.

d. Generalized Norm and Measurement Noise

In this section the results of Sections II.6.b and II.6.c are considered for the minimization of the more general output error functional

$$J_Q(\hat{b}) = [(y - y(\hat{b}))/Q(y - y(\hat{b}))]$$

$$= \int_{t_o}^{t_f} [y(t) - y(t;\hat{b})/Q(t)(y(t) - y(t;\hat{b}))]\, dt \tag{51}$$

For each $t \in [t_o, t_f]$ the m x m matrix $Q(t) \in R^{m \times m}$ is assumed to be symmetric and uniformly bounded and positive definite[1]. Such an output error functional would generally find application in maximum likelihood identification when there is zero-mean additive white Gaussian measurement noise, $\eta \in Y$, with covariance

---

[1]That is, there exists positive constants, c_1 and c_2, such that

$$c_1[\cdot/\cdot] \leq [\cdot/Q(t)\cdot] \leq c_2[\cdot/\cdot] \quad \text{for all } t \in [t_o, t_f].$$

$$E(\eta(t)\eta^T(\tau)) = Q^{-1}(t)\delta(t - \tau) \tag{52}$$

(see, e.g., (Ref 85)). In the notation above, $E(\cdot)$ is the expectation operator, the superscript "T" is the transpose, and $\delta(\cdot)$ is the unit sampling function defined by $\delta(t) = 0$ for $t \neq 0$, and

$$\int_{-\varepsilon}^{\varepsilon} \delta(t)g(t)dt = g(0) \tag{53}$$

for any arbitrary function g.

Using the assumptions concerning Q and introducing the Q-inner product defined by[1]:

$$[y/y]_Q = [y/Qy] \tag{54}$$

for $y \in \mathcal{Y} = L_2(t_o, t_f; R^m)$, it is straight-forward to show that the Hilbert space $\{\mathcal{Y}, [\cdot/\cdot]\}$ is isometrically isomorphic to the Q-Hilbert space $\{\mathcal{Y}, [\cdot/\cdot]_Q\}$ (Ref 27: 84 - 85), and so the two may be treated interchangeably. In terms of the Q-Hilbert space, the output error functional $J_Q(\hat{b})$ takes the form

$$J_Q(\hat{b}) = [y - y(\hat{b})/y - y(\hat{b})]_Q \tag{55}$$

Thus, our previous results concerning $J_e(\hat{b})$ can be readily transferred to this more general case.

To illustrate, consider the adjoint of V mapping from the Hilbert space $\{R^p, [\cdot/\cdot]\}$ into the Hilbert space $\{\mathcal{Y}, [\cdot/\cdot]_Q\}$. Suppose that V^* is defined by

$$[y/Vb] = [V^*y/b] \tag{56}$$

for $y \in \mathcal{Y}$, $b \in R^p$. Then

---

[1]Inner products not subscripted are assumed to be the natural inner product.

$$[y/Vb]_Q = [y/QVb] = [V^*Qy/b]. \tag{57}$$

Therefore, V^*Q is the adjoint of V mapping between $\{R^p,[\cdot/\cdot]\}$ and $\{\mathcal{Y},[\cdot/\cdot]_Q\}$. Thus, while the matrix V^*V was of fundamental importance in minimizing $J_e(\hat{b})$, we see that the counterpart matrix

$$V^*QV = \int_{t_o}^{t_f} [v^{(i)}(t)/Q(t)v^{(j)}(t)]dt \tag{58}$$

is basic to minimizing $J_Q(\hat{b})$. If Q has the physical meaning described in the first paragraph of this section, then this matrix is termed the "information matrix" (Ref 31). It has general importance in estimation theory because the covariance of estimation error for any unbiased estimator is limited by the Cramer-Rao lower bound (Ref 31):

$$E((b - \hat{b}^*)(b - \hat{b}^*)^T) \geq (V^*(b)QV(b))^{-1} \tag{59}$$

where $b \in R^p$ is the true parameter value.

With this short introduction into the relationship between parameter sensitivity and system identification, we now discuss the related topics of identifiability and sensitivity design for optimal identification.

7. SYSTEM PARAMETER IDENTIFIABILITY

In this section we discuss parameter identifiability. Background literature on this topic was presented in Section I.1.d. Like the sections on insensitivity and controllability our primary purpose here is to establish the definitions and an operator methodology for treating identifiability. In Section IV this approach will be applied to linear time-invariant ordinary differential equation systems to yield algebraic conditions for identifiability.

Roughly stated, parameter identifiability is the question of determining a priori system conditions which ensure that a unique best

estimate of the unknown parameters may be obtained from the observable output. Such a concept of identifiability was introduced by Lee (Ref 19). As noted in Section I.1.d various other probabilistic interpretations of identifiability have been utilized; however, uniqueness of solution is most consistent with our deterministic approach to parameter identification, and it is the concept of identifiability which has been used by quite a number of other researchers in the determination of structural properties of identifiability. (See Section I.1.d for discussion.)

To motivate the definition of local identifiability which we adopt, again consider the linearized measurement equation

$$z = y - y(b_o) = V\Delta b + \varepsilon_1 + \eta \tag{60}$$

The bias ε_1 is of second order in Δb and the measurement noise process $\eta \in \mathcal{Y}$ is immaterial for our deterministic, "least squares" approach. (See Section II.6.d.) Then for values of b which are local to b_o equation (34) becomes approximately

$$z = V\Delta b. \tag{61}$$

For a given measurement error process $z \in \mathcal{Y}$, it is well-known that the above equation has a unique solution for Δb if and only if the operator V is one-to-one. Therefore, we define:

Definition 5

The system S_H is locally identifiable at $b_o \in R^p$ if $V \in L_c(R^p, \mathcal{Y})$ is one-to-one.

Definition 6

The locally non-identifiable subspace at $b_o \in R^p$ is the null space of V while the locally identifiable subspace is the orthogonal complement of the non-identifiable subspace $M(b_o) = N(V)^{\perp}$.

The non-identifiable subspace is closely related to the concepts of insensitivity. In words, the non-identifiable subspace is the subspace in which parameter perturbations will cause no first order change in the observable output. This relationship between system parameter insensitivity and system non-identifiability has been recognized previously and is discussed further in Section IV. (See, for example, Bonivento (Ref 55).)

V is bounded linear operator from the finite dimensional space R^p to the function space Y and so must have a finite dimensional mange space. Thus $R(V)$ must be closed in Y, and so we may immediately obtain the following lemma:

Lemma 4

The system S_H is locally identifiable at $b_o \in R^p$ if and only if the p x p dimensioned symmetric "identifiability matrix"

$$M(b_o) = V^*V \tag{62}$$

has an inverse. Further, the non-identifiable subspace is

$$N(M(b_o)) \subset R^p \tag{63}$$

while the identifiable subspace is

$$M(b_o) \equiv R(M(b_o)) \subset R^p \tag{64}$$

Proof

The first two parts of the theorem follow directly from Definitions 5 and 6 and standard results from functional analysis (see, e.g., Taylor (Ref 27: 250 - 251)). The last part of the theorem follows from the fact that $M(b_o)$ is a symmetric matrix and property 2 of Lemma 1.

Q.E.D.

From Lemma 4 and the results of Section II.6.b (see equation (36)) we see that system S_H is locally identifiable if and only if we may obtain a unique solution of linearized measurement equation error functional, $J_{LM}(\Delta b)$, equation (35). Thus we see the relationship between the definition for local identifiability and the ability to obtain, locally, a unique minimum norm estimate of b. Indeed, Lemma II.5 could be taken as the definition of local identifiability and definitions 5 and 6 then derived as lemmas to this alternate definition. This is the approach which several others have taken (e.g., Martenson (Ref 96), Glover and Willems (Ref 93)), but either approach is equivalent. Also, in the comments of Section II.6.c we noted that the local information matrix must be positive definite to ensure the existence of local steps in the steepest descent and Newton-Raphson iterations, as well.

Since the self-adjoint operator Q of our generalized norm functional, equation (51), is assumed invertible, it is easy to show that

a. $R(V^*QV) = R(V^*V)$ (65)

b. $N(V^*QV) = N(V^*V)$ (66)

c. $(V^*QV)^{-1}$ exists if and only if $(V^*V)^{-1}$ exists.

Thus the local information matrix, $M_Q(b_o) = V^*QV$, may be substituted for the identifiability matrix, $M(b_o)$, in Lemma 4, and the conclusions remain the same. With this substitution, Lemma 4 is a restatement of the well-known result that the system is locally identifiable if and only if the local information matrix, $M_Q(b_o)$, is positive definite. (See, for example, Tse (Ref 102).)

In Section IV we apply these results to obtain local identifiability conditions for linear ordinary differential equation systems. For time

varying systems, Lemma 4 and the "identifiability matrix" may be used to test for local identifiability; however, for time-invariant systems an algebraic description of the sensitivity operators allows us to use Definitions 5 and 6 directly and thereby obtain algebraic criteria for local identifiability. Using these algebraic conditions we will see that we may not only test for local identifiability at the nominal $b_o \in R^p$, but we may treat $b \in R^p$ as a variable and test for identifiability as a function of b. Such a concept was introduced by Gupta and Hall (Ref 108) in consideration of "structural" controllability of the sensitivity operators of linear time-invariant ordinary differential equation systems. Therefore, we define:

Definition 7

The system S_H is structurally identifiable if $V(b) \in L_c(R^p, Y)$ is one-to-one for almost all $b \in R^p$.

Such a structural concept is particularly significant for identifiability because one wishes to ensure that the system will be "locally" identifiable along the entire sequence of possible parameter estimates.

Finally not only may it be important to know whether the parameters of a linear system are locally identifiable from the total system output, but it may be equally important to know whether the parameters are locally identifiable from the zero-input response alone or from the zero-state response alone or from both. For example, such information might be important in the design of experiments for parameter identification (see Section II.8). This information is easily obtained by considering whether or not the zero-input and zero-state sensitivity operators are one-to-one, respectively. Therefore, definitions parallel to Definitions 5 and 6 can be made for zero-input and zero-state local

identifiability, and Lemma 4 may be given appropriate corollaries for the "zero-input" and "zero-state" local identifiability matrices. Also, just as the case for parameter insensitivity, we will find that it is more convenient to obtain separate conditions for zero-input and zero-state local identifiability rather than obtaining joint conditions (see Section IV). The following lemma relates the separate zero-input and zero-state locally identifiable subspaces to the total locally identifiable subspace:

Lemma 5

Let $V_{z.i.}$ and $V_{z.s.}$ be the zero-input and zero-state sensitivity operators for the system S_H evaluated at the nominal $b_o \in R^p$. Suppose that $V_{z.i.}\, b \neq -V_{z.s.}\, b$ for all

$$b \in (N(V_{z.i.}))^{\perp} \subset R^p,\ b \neq 0. \tag{67}$$

Then the total system non-identifiable subspace is

$$N(V) = N(V_{z.i.}) \cap N(V_{z.s.}) \tag{68}$$

and the total identifiable subspace is

$$M(b_o) = M_{z.i.}(b_o) + M_{z.s.}(b_o) \tag{69}$$

Proof

Since $V = V_{z.i.} + V_{z.s.}$ it is clear that

$$N(V_{z.i.}) \cap N(V_{z.s.}) \subset N(V). \tag{70}$$

Suppose $b \in N(V)$ and $b \neq 0$. Then

$$V_{z.i.}\, b + V_{z.s.}\, b = 0, \tag{71}$$

and so either $b \in N(V_{z.i.})$ or $b \in (N(V_{z.i.}))^{\perp}$. But if $b \in (N(V_{z.i.}))^{\perp}$ this would contradict the assumption (67). Hence $b \in N(V_{z.i.})$. This along with equation (71) implies that $b \in N(V_{z.s.})$. Thus

$$N(V) \subset N(V_{z.i.}) \cap N(V_{z.s.}) \tag{72}$$

and the assertion (68) is proved.

Finally, since $N(V) \subset R^P$, equation (69) is obtained by taking the orthogonal complement of both sides of equation (68). (See, for example, Nering, Theorem 4.4 (Ref 21: 140).)

Q.E.D.

In this section we have considered the binary question of whether or not the parameters of a linear system are locally identifiable at an a priori parameter value $b_o \in R^P$. We have not discussed the "quality" of the identification (estimation) capability. This important experimental design question is discussed in the next section.

8. SENSITIVITY OPERATOR DESIGN FOR OPTIMAL PARAMETER IDENTIFICATION

In Section II.6 we discussed the fact that the "quality" of parameter identification capability is strongly dependent upon the parameter sensitivity operators. Heuristically, if we consider the linearized measurement equation (34) of Section II.6.a with process noise $\eta \in Y$ and a priori parameter value $b_o \in R^P$,

$$z = V\Delta b + \varepsilon_1 + \eta \tag{34}$$

then the sensitivity operator V may be likened to the signal strength in a communication channel. For improved identification, one would like to maximize the "signal-to-noise" ratio. If Q^{-1} is the covariance matrix of the zero-mean Gaussian white noise process (see Section II.6.d)

then the local information matrix, $M_Q = V^*QV$ may be viewed as the "signal-to-noise" ratio for the multi-channel, multi-parameter, estimation problem, equation (34). Also, we commented in Section II.6.d the information matrix evaluated at the true system parameter value, b, provides a lower bound for the estimation error covariance matrix of any unbiased estimate through the Cramer-Rao inequality (59). Since we do not know the true parameter value, b, we can only approximate this lower bound through the use of the local information matrix, M_Q.

Therefore, based upon the preceding discussion, we will use some measure on the local information matrix as a criterion of optimality to "design" system nominal sensitivity operators for improved parameter identification. Indeed, for just these same reasons the local information matrix has been used as a design criterion for improved parameter identification by quite a number of previous researchers (see Section I.1.e for a discussion of the relevant literature on input design). Since an optimization criterion generally requires a scalar quantity, some suitable scalar measure on the p x p information matrix must be obtained. This question is discussed briefly in the next subsection.

If we assume that both the initial condition vector, $d \in D$, and the control input $u \in U$, are independent of the unknown parameter vector, $b \in R^p$, and that they may be selected arbitrarily, then the nominal sensitivity operators are functions of both of these quantities. To be more explicit, the zero-input sensitivities are linear operators on $d \in D$

$$v^{(i)}_{z.i.}(d) = (T)_{(i)}d \tag{73}$$

while the zero-state sensitivities are linear operators on the control $u \in U$

$$v^{(i)}_{z.s.}(u) = (W)_{(i)}u \tag{74}$$

i = 1, 2, ...p. Therefore, both d and u may be selected to optimize some measure of the local information matrix. This optimization could be accomplished concurrently; however, it is computationally more convenient to consider the separate selection of the initial condition vector which optimizes the "zero-input" local information matrix on the one hand and the control function which optimizes the "zero-state" local information matrix on the other. These computational advantages will be understood more clearly when this topic is discussed for linear time-invariant ordinary differential equation systems in Section VII. Also, this separate design of the initial condition vector and the control input based upon the zero-input and zero-state responses, respectively, is related to the previous section concerning the separate determination of the zero-input and zero-state identifiable subspaces. Indeed, certain parameter components may be identifiable from one response while not from the other. Thus, from a practical point-of-view, it may be desirable to design separate experiments with zero initial conditions in one case and zero control inputs in the other.

Finally, we comment that the approach taken in this section is motivated by the work of Mehra (Ref 112). Mehra considers the problem of control input design for optimal parameter identification in linear time-invariant ordinary differential equation systems. For the special case of zero initial conditions, he obtains an operator formulation of the optimization problem and shows that the optimal control is the eigenfunction corresponding to the maximal eigenvalue of a positive, self-adjoint operator. The approach taken here is a generalization of Mehra's work, and conceptually there is very little difference. However, it should be mentioned that the concept of formally optimizing the initial

condition vector to provide optimal parameter identification capability in the transient response is a new concept to the current theoretical literature on parameter identification. On the other hand, the manipulation of the initial condition vector to provide improved identification is a well-known, and often very important, ad-hoc procedure for providing parameter identification capability in practical applications. (See, for example, (Ref 85) in which initial platform orientation plays an essential role in the identification of inertial platform parameters.) The capability to determine optimal initial conditions for identification purposes might be particularly important in cases in which it is either impossible or very costly to provide a control input in order to identify system parameters.

a. Optimization Criterion

Because the information matrix is dimension p x p, some suitable scalar measure must be chosen as an optimization criterion. This problem has been discussed by a number of researchers, and Section I.1.e reviews the literature on this subject.

From an information theory viewpoint, either a linear functional of the dispersion matrix (inverse of the local information matrix) or the determinant of the dispersion matrix are the most appropriate criterion. From the Cramer-Rao lower bound we know that the dispersion matrix is a lower bound for the covariance matrix of the parameter estimation error, and so a linear functional of the dispersion matrix (usually a weighted trace) is directly related to the variance of estimation error. On the other hand, the determinant of the dispersion matrix is physically appealing because it is proportional to the volume of the error-ellipsoid in p-dimensional parameter space.

However, the direct use of either of these criteria leads to difficult computational problems in an optimization algorithm. Therefore, a criterion used most frequently because of the convenient quadratic functional which it provides is a linear functional on the information matrix itself. For this reason we choose a weighted trace of the information matrix

$$J_K(d,u) = \sum_{i,j=1}^{p} k_{ij} [v^{(i)}/Qv^{(j)}] \tag{75}$$

as our optimization criterion. It is fairly well-known (e.g. (Ref 114), (Ref 119), and (Ref 122)) that such a criterion can lead to identifiability problems and poor estimates of the parameters if the weighting constants k_{ij} are not selected judiciously. However, Nahi and Napjus (Ref 117), Mehra (Ref 114) and (Ref 115), and Gupta and Hall (Ref 108) develop iterative algorithms to adjust the weightings k_{ij}, i, j=1, 2, ...p, so that the resulting maximum of J_K will minimize either a linear functional on the dispersion matrix or the determinant of the dispersion matrix. The maximization of J_K is a basic and most time-consuming part of each of these iterative techniques, and so it is quite apropos to utilize J_K as our optimization criterion.

b. Selection of Initial Conditions

In this section we assume that the system S_H has zero control input and that we wish to select an initial condition vector, $d^* \in D$, which optimizes parameter identification capability from the zero-input response. The optimization functional takes the form

$$J_{K_{z.i.}}(d) = \sum_{i,j=1}^{p} k_{ij}[v^{(i)}_{z.i.}(d)/Qv^{(j)}_{z.i.}(d)]$$

$$= \sum_{i,j=1}^{p} k_{ij}[T_{(i)}d/QT_{(j)}d] \tag{76}$$

If we let $T^*_{(i)} \in L_c(Y,D)$ denote the adjoint of $T_{(i)} \in L_c(D,Y)$ then $J_{K_{z.i.}}(d)$ takes the form

$$J_{K_{z.i.}}(d) = [d/A_{K_{z.i.}}d] \tag{77}$$

where $A_{K_{z.i.}} \in L_c(D,D)$ is the positive, bounded self-adjoint operator

$$A_{K_{z.i.}} = \sum_{i,j=1}^{p} k_{ij}\, T^*_{(i)}QT_{(j)}. \tag{78}$$

For linear ordinary differential equation systems, D is the finite dimensional space R^n, and so $A_{K_{z.i.}}$ is a positive, symmetric n x n matrix.

Thus $J_{K_{z.i.}}(d)$ is a quadratic form in d. Now in order to obtain a unique maximum of this functional, some auxiliary constraint must be placed upon the initial condition vector, $d \in D$. It is convenient to assume that d must satisfy the "energy" constraint[1]

$$[d/d] \leq 1. \tag{79}$$

---

[1]Note that there is no loss in generality by this assumption; for, suppose we wish to satisfy the more general constraint $[d^*/Ed^*] \leq 1$, where E is a strictly positive, self-adjoint operator. Defining the E-inner product

$$[\cdot/\cdot]_E = [\cdot/E\cdot]$$

it is easily shown that the Hilbert space $\{D,[\cdot/\cdot]_E\}$ is isometrically isomorphic to the natural Hilbert space $\{D,[\cdot/\cdot]\}$. Therefore, the constraint which we have assumed may easily be extended to this more general case.

Then it is well known (e.g., Blum (Ref 3: 601)) that the $d^* \in D$ which maximizes $J_{K_{z.i.}}(d)$ and satisfies the constraint (79), is the eigenvector of $A_{K_{z.i.}}$ which corresponds to the largest eigenvalue of the self-adjoint operator $A_{K_{z.i.}}$ In such a case, $[d^*/d^*] = 1$, and the eigenvector need not be unique. If the operator $A_{K_{z.i.}} \in L_c(D,D)$ is compact (which it will be if $D = R^n$), then we are assured that a maximizing eigenvector, $d^* \in D$, does exist. (See, for example, Blum (Ref 3: 601).)

In Section VII we consider this problem formulation for linear time-invariant ordinary differential equation systems, and the positive, symmetric n x n matrix, $A_{K_{z.i.}}$ is easily computed. Obtaining the optimal initial condition vector is then merely a matter of determining the maximizing eigenvector of $A_{K_{z.i.}}$.

c. Selection of Control Input

In this section we assume that the system S_H has zero initial conditions and that we wish to select a control input, $u^* \in U$, which optimizes parameter identification capability from the zero-state response. Thus the optimization functional takes the form

$$J_{K_{z.s.}}(u) = \sum_{i,j=1}^{p} k_{ij}[v^{(i)}_{z.s.}(u)/Qv^{(j)}_{z.s.}(u)]$$

$$= \sum_{i,j=1}^{p} k_{ij}[W_{(i)}u/QW_{(j)}u] \qquad (80)$$

This problem is a complete parallel of the zero-input design problem of the previous section. Let $W^*_{(i)} \in L_c(Y,U)$ denote the adjoint operator of $W_{(i)} \in L_c(U,Y)$. Then

$$J_{K_{z.s.}}(u) = [u/A_{D_{z.s.}}u] \qquad (81)$$

where $A_{K_{z.s.}} \in L_c(U,U)$ is the positive, bounded, self-adjoint operator

$$A_{K_{z.s.}} = \sum_{i,j=1}^{p} k_{ij} W^*_{(i)} Q W_{(j)} \tag{82}$$

Again we assume a control "energy" constraint[1]

$$[u/u] \leq 1. \tag{83}$$

Then the $u^* \in U$ which maximizes $J_{K_{z.s.}}(u)$ and satisfies the constraint (83), is the eigenfunction corresponding to the largest eigenvalue of the positive, self-adjoint operator $A_{K_{z.s.}}$ Again, it will be true that $[u^*/u^*] = 1$, and the maximizing eigenfunction need not be unique. Also, if $A_{K_{z.s.}}$ is compact, we are assured the existence of the maximizing eigenfunction. (See, Blum, (Ref 3: 601).) The operator $A_{K_{z.s.}}$, equation (82), will be compact if each of the partials $W_{(i)}$, $i = 1, 2, \ldots p$, are compact (products of bounded linear operators and compact operators are compact, and linear combinations of compact operators are compact (Ref 27: 274)). For our later application to linear ordinary differential equation systems, the operators $W_{(i)}$ are Fredholm integral equations on the compact interval $[t_o, t_f]$ and have continuous kernels, and so they will indeed be compact (see Liusternik and Sobolev (Ref 19: 129)).

The above presentation is a generalization of the results of Mehra (Ref 112) to control input design for the Hilbert space system S_H. Mehra suggests various computational methods of solution for this problem, but

---

[1]Note that there is no loss in generality by this assumption; for, suppose we wish to satisfy the more general constraint $[u^*/Eu^*] \leq 1$, where E is a strictly positive, self-adjoint operator. Defining the E-inner product

$$[\cdot/\cdot]_E = [\cdot/E\cdot]$$

it is easily shown that the Hilbert space $\{U,[\cdot/\cdot]_E\}$ is isometrically isomorphic to the natural Hilbert space $\{U,[\cdot/\cdot]\}$. Therefore, the constraint which we have assumed may easily be extended to this more general case.

we will not discuss these until Section VII. There we will examine computational techniques associated with input design for linear time-invariant ordinary differential equation systems.

9. SUMMARY

This concludes our treatment of parameter sensitivities for the Hilbert space linear system, S_H. Some known concepts have been extended to this Hilbert space setting (e.g., parameter insensitivity, sensitivity controllability, parameter identification techniques, local parameter identifiability, and input design for improved identification), and some new techniques have been suggested (e.g., an operator approacn to minimum sensitivity control design and selection of an initial condition vector to optimize identification capability from the zero-input response). But the primary contribution of this section has been to present a unified operator approach for treating parameter sensitivity and various sensitivity-related system properties.

This section lays the groundwork for our application to linear ordinary differential equation systems considered in the subsequent sections. We concentrate on time-invariant systems because it is here that a new algebraic description of the parameter sensitivities provides convenient geometric and computational tools for treating the system properties. This algebraic description of the parameter sensitivity operators is developed in the next section.

Section III

SENSITIVITY OPERATORS FOR LINEAR ORDINARY DIFFERENTIAL EQUATION SYSTEMS

In this section operator expressions are derived for parameter sensitivity of continuous linear ordinary differential equation systems. Both time-varying and time-invariant systems are considered. These results illustrate the operator representation of the parameter sensitivities which was presented in Section II for the Hilbert space linear system S_H. Furthermore, these results form the foundation for our subsequent discussion of the sensitivity-related system properties in Sections IV - VII.

The organization of this section is the following: in Section III.1 operator expressions are derived for the parameter sensitivities in time-varying linear systems. The key to this operator representation is to obtain the partial derivatives of the state transition matrix with respect to the parameter components. A general theorem concerning these partial derivatives is proved; however, for systems with a time-varying plant matrix the computational aspects of the operator representation are less desirable than those for computing the sensitivities from sensitivity system differential equation techniques. In Section III.2 operator expressions are derived for the parameter sensitivities in linear time-invariant systems. Again the key factor in this representation is taking the partial derivatives of the state transition matrix, e^{At}. General polynomial expressions for these partial derivatives are obtained and useful relationships to eigenvalue sensitivities are shown.

This polynomial description of the partial derivatives of e^{At} is used to provide a fundamental matrix-operator representation of the

parameter sensitivities in linear time-invariant systems. This form proves to be computationally useful and also provides considerable geometric insight for the study of various sensitivity related system properties. These computational and geometric features are examined in Sections IV - VII.

1. TIME-VARYING SYSTEMS

In this section we derive operator expressions for the parameter sensitivities of the time-varying linear system, S_{T_V}:

$$\dot{x}(t;u) = A(t)x(t;u) + B(t)u(t) \tag{84}$$

$$x(t_o) = d \qquad t \in [t_o, t_f] \tag{85}$$

$$x(t;u) \in R^n \qquad u(\cdot) \in L_2(t_o, t_f; R^n) \tag{86}$$

with observable output

$$y(t;u) = C(t)x(t;u) \tag{87}$$

$$y(t;u) \in R^m \qquad t \in [t_o, t_f] \tag{88}$$

where the matrices A, B, C and the initial condition vector d are assumed to be real and parameterized by the real p-dimensioned parameter vector b with nominal value $b_o \in R^p$. Again we will observe the convention of not showing the explicit dependency upon b unless it is required for clarity and it will always be assumed that quantities are evaluated at b_o unless indicated otherwise. For all $t \in [t_o, t_f]$ and at b_o it is assumed that A(t), B(t), C(t), and d, are each continuously differentiable with respect to every parameter component, b_i, i = 1, 2, ...p. Furthermore, for all $b \in R^p$ within some neighborhood of $b_o \in R^p$, it is assumed that

A(.;b), B(.;b), and C(.;b) and all p of their partial derivatives are uniformly bounded and piecewise continuous on the finite interval $[t_o, t_f]$.

Under these assumptions (Ref 29) the sensitivities satisfy the following linear "sensitivity system":

$$\dot{X}(t;u) = \bar{A}(t)X(t;u) + \bar{B}(t)u(t) \tag{89}$$

$$X(t_o;u) = \bar{d} \qquad t \in [t_o, t_f] \tag{90}$$

$$Y(t;u) = \bar{C}(t)X(t;u) \tag{91}$$

where

$$X(t;u) = \begin{bmatrix} x(t;u) \\ \xi^{(1)}(t;u) \\ \vdots \\ \xi^{(p)}(t;u) \end{bmatrix} \qquad Y(t;u) = \begin{bmatrix} y(t;u) \\ v^{(1)}(t;u) \\ \vdots \\ v^{(p)}(t;u) \end{bmatrix} \tag{92}$$

$$\bar{A}(t) = \begin{bmatrix} A(t) & 0 & \cdots & 0 \\ A_{(1)}(t) & A(t) & & \\ \vdots & \vdots & & \vdots \\ A_{(p)}(t) & 0 & & A(t) \end{bmatrix} \qquad \bar{B}(t) = \begin{bmatrix} B(t) \\ B_{(1)}(t) \\ \vdots \\ B_{(p)}(t) \end{bmatrix} \tag{93}$$

$$\bar{C}(t) = \begin{bmatrix} C(t) & 0 & \cdots & 0 \\ C_{(1)}(t) & C(t) & & \\ \vdots & \vdots & & \vdots \\ C_{(p)}(t) & 0 & \cdots & C(t) \end{bmatrix} \qquad \bar{d} = \begin{bmatrix} d \\ d_{(1)} \\ \vdots \\ d_{(p)} \end{bmatrix} \tag{94}$$

and $\xi^{(i)}(t;u)$ and $v^{(i)}(t;u)$, $i = 1, 2, \ldots p$, are the state and output sensitivities respectively. Therefore, the solution of the sensitivity system requires solving $n(p + 1)$ coupled linear differential equations.

However, we wish to obtain an operator solution of the parameter sensitivities as described in Section II.1. Therefore, consider the operator solution of the system S_{TV}:

$$x(t;u) = \Phi(t,t_o)d + \int_{t_o}^{t} \Phi(t,s)B(s)u(s)ds \tag{95}$$

$$\begin{aligned} y(t;u) &= C(t)\, x(t;u) \\ &= T(t)d + W(t)u \end{aligned} \tag{96}$$

where $\Phi(t,t_o)$ is the state transition matrix of $A(\cdot)$ and the zero-input and zero-state operators $T(t) \in L_c(R^n,R^m)$ and $W(t) \in L_2(L_c(t_o,t;R^r),R^m)$ are defined by

$$T(t) = C(t)\Phi(t,t_o) \tag{97}$$

$$W(t) = \int_{t_o}^{t} C(t)\Phi(t,s)B(s)(\cdot)ds \tag{98}$$

respectively. For discussion we will deal almost exclusively with the output sensitivity operators, and these are given by the operator expressions

$$v^{(i)}_{z.i.}(t) = T_{(i)}(t)d + T(t)d_{(i)} \tag{99}$$

$$v^{(i)}_{z.s.}(t;u) = W_{(i)}(t)u \tag{100}$$

To compute the partial derivatives of $T(t)$ and $W(t)$ requires that the partial derivatives of the state transition matrix, $\Phi(t,s)$, be obtained. The next theorem establishes the existence and continuity of

these partial derivatives. We comment that this theorem is a well-known result for the case in which $A(\cdot)$ is time-invariant (Bellman (Ref 2)); however, our proof for this time-varying case is different than the one used by Bellman, and so the proof is instructive in its own right.

Theorem 1

Suppose for all $t \in [t_o, t_f]$ at the nominal parameter value, $b_o \in R^p$, $A(t) \in R^{nxn}$ is continously differentiable with respect to each parameter component, b_i, $i = 1, 2, \ldots p$; and for all $b \in R^p$ in a neighborhood of b_o, $A(t;b)$ and all p of its partial derivatives are uniformly bounded and piecewise continuous in t. Next, suppose $\Phi(t,s)$ is the state transition matrix of $A(t)$; that is, for all $t, s \in [t_o, t_f]$ $\Phi(t,s)$ is the unique solution of the matrix differential equation

$$\frac{d}{dt}\Phi(t,s) = A(t)\Phi(t,s) \tag{101}$$

$$\Phi(s,s) = I \tag{102}$$

Then for each $t, s \in [t_o, t_f]$ the partial derivatives $\Phi_{(i)}(t,s)$, $i = 1, 2, \ldots p$, may be obtained either from the integral equation

$$\Phi_{(i)}(t,s) = \Phi(t,s) \int_s^t \Phi(s,\zeta)A_{(i)}(\zeta)\Phi(\zeta,s)d\zeta \tag{103}$$

or the matrix differential equation

$$\frac{d}{dt}\Phi_{(i)}(t,s) = A(t)\Phi_{(i)}(t,s) + A_{(i)}(t)\Phi(t,s) \tag{104}$$

with boundary condition $\Phi_{(i)}(s,s) = 0$. Furthermore, these partial derivatives are continuous in b.

Proof

Let $A_{\Delta_i}(t)$ denote $A(t;b_o + \Delta b_i e_i)$ where e_i is the i^{th} Euclidean basis vector and $\Phi_{\Delta_i}(t,s)$ denote the unique solution of the linear matrix differential equation

$$\frac{d}{dt}\Phi_{\Delta_i}(t,s) = A_{\Delta_i}(t)\Phi_{\Delta_i}(t,s) \tag{105}$$

$$\Phi_{\Delta_i}(s,s) = I \tag{106}$$

Assuming that $b_o + \Delta b_i e_i$ is contained in the neighborhood in which $A(t;b)$ is uniformly bounded and piecewise continuous, then both $\Phi(t,s)$ and $\Phi_{\Delta_i}(t,s)$ will be uniformly bounded for all $t, s \in [t_o, t_f]$. The differential equation for $\Phi_{\Delta_i}(t,s)$ may be written as

$$\frac{d}{dt}\Phi_{\Delta_i}(t,s) = A(t)\Phi_{\Delta_i}(t,s) + (A_{\Delta_i}(t) - A(t))\Phi_{\Delta_i}(t,s) \tag{107}$$

which has the unique solution

$$\Phi_{\Delta_i}(t,s) = \Phi(t,s) + \Phi(t,s)\int_s^t \Phi(s,\zeta)(A_{\Delta_i}(\zeta) - A(\zeta))\Phi_{\Delta_i}(\zeta,s)d\zeta \tag{108}$$

Using the Bounded Convergence Theorem (Ref 25: 81) and the fact that $A(.;b)$ is continuous in b at b_o, the limit of both sides of equation (108) gives

$$\lim_{|\Delta\theta_i| \to 0} \Phi_{\Delta_i}(t,s) = \Phi(t,s) \tag{109}$$

Hence, $\Phi(t,s)$ is also continuous in b at b_o.

Since the partial derivatives of $A(t;b)$ are uniformly bounded in a neighborhood by b_o, we may use equation (106) and again apply the Bounded Convergence Theorem to the definition of the partial derivative to obtain

$$\Phi_{(i)}(t,s) \equiv \lim_{|\Delta b_i| \to 0} \frac{\Phi_{\Delta_i}(t,s) - \Phi(t,s)}{\Delta b_i}$$

$$= \Phi(t,s)\int_s^t \Phi(s,\zeta)A_{(i)}(\zeta)\Phi(\zeta,s)d\zeta \tag{110}$$

Since $A_{(i)}(\cdot)$ is assumed continuous in b, all the factors on the right hand side of equation (110) are continuous in b. Thus $\Phi_{(i)}(t,s)$ is also continuous in b at b_o.

Finally, the differential equation representation of $\Phi_{(i)}(t,s)$ follows by application of Leibnitz Rule to equation (103).

Q.E.D.

Because of the uniform bounds assumed on the partial derivatives of $A(\cdot)$ and $B(\cdot)$, the partial derivative may also be brought inside the integral in equation (98). Therefore, the partial derivatives of the operators T(t) and W(t), equations (97) and (98), may be described in terms of the partial derivatives of C(t), B(t), and $\Phi(t,t_o)$ where the latter partial derivative is computed from Theorem 1. These may then be used in equations (99) and (100) to give an operator expression for the zero-input and zero-state sensitivities. However, because of the convolution in equation (98), this is not a particularly practical way for computing the parameter sensitivities for this time-varying plant matrix case. For the time-invariant system considered in the next section, these undesirable convolutions may be transformed into computationally efficient quadratures.

2. TIME-INVARIANT SYSTEMS

In this section we assume that the A, B, and C matrices in equations (84) and (87) are time-invariant. Also, without loss in generality we will assume that $t_o = 0$. This time-invariant ordinary differential equation system will be denoted S_{LC}. Like the time-varying system of the previous subsection, the key to obtaining a practical operator form of the parameter sensitivities will be an efficient way to compute the partial derivatives of the state transition matrix, $\Phi(t) = e^{At}$. Theorem 1

could be utilized; however, for the time-invariant A matrix we are able to obtain far more fundamental and useful forms of these partial derivatives.

a. Mathematical Description of $e^{At}_{(i)} = \Phi_{(i)}(t)$

Consider the system matrix $A \in R^{nxn}$ at the a priori parameter value $b_o \in R^p$. The following notation will be used throughout this section. Suppose that A has characteristic polynomial

$$\Delta(q) = \det(qI - A) = \prod_{k=1}^{\rho} (q - q_k)^{n_k} \tag{111}$$

and μ^{th} order minimal polynomial

$$\psi(q) = \prod_{k=1}^{\rho} (q - q_k)^{\mu_k} \tag{112}$$

where q_k, $k = 1, 2, \ldots\rho$, are the distinct eigenvalues of A; $n_k \geq \mu_k \geq 1$ are the multiplicities of q_k in the characteristic and minimal polynomials respectively; and $\mu = \sum_{k=1}^{\rho} \mu_k$ and $n = \sum_{k=1}^{\rho} n_k$.

The system state transition matrix may then be represented as

$$\Phi(t) = e^{At} = \sum_{k=1}^{\rho} \sum_{j=0}^{\mu_k - 1} Z_{k,j} t^j e^{q_k t} \tag{113}$$

where the n x n matrices $Z_{k,j}$ are termed the components of A. (Ref 33: 604). An equivalent representation is (Ref 33: 609)

$$e^{At} = \sum_{j=1}^{\mu} A^{j-1} \alpha_j(t) \tag{114}$$

where the real, scalar, linearly independent functions, $\alpha_j(\cdot)$, $j = 1, 2, \ldots\mu$, are uniquely determined from the μ linear equations

$$t^j e^{q_k t} = \frac{d^j}{dq^j} [1 \; q \; q^2 \; \ldots \; q^{\mu-1}]_{q=q_k} \begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \vdots \\ \alpha_\mu(t) \end{bmatrix} \tag{115}$$

$k = 1, 2, \ldots \rho$

$j = 0, 1, \ldots \mu_k - 1$

We shall presently determine representations for the partial derivative $e^{At}_{(i)}$ which are equivalent to each of the above forms.

Now consider the i^{th} augmented "sensitivity system":

$$\begin{bmatrix} \dot{x}(t;u) \\ \dot{\xi}^{(i)}(t;u) \end{bmatrix} = \begin{bmatrix} A & 0 \\ A_{(i)} & A \end{bmatrix} \begin{bmatrix} x(t;u) \\ \xi^{(i)}(t;u) \end{bmatrix} + \begin{bmatrix} B \\ B_{(i)} \end{bmatrix} u(t) \tag{116}$$

$$\begin{bmatrix} x(0) \\ \xi^{(i)}(0) \end{bmatrix} = \begin{bmatrix} d \\ d_{(i)} \end{bmatrix} \tag{117}$$

The augmented matrix

$$\tilde{A}_i \equiv \begin{bmatrix} A & 0 \\ A_{(i)} & A \end{bmatrix} \tag{118}$$

is of fundamental importance in describing the operator solution of this i^{th} sensitivity system, (116), and we shall obtain several useful properties concerning parameter sensitivity, in general, by consideration of this augmented matrix. Indeed, the first theorem provides the foundation for all of our remaining results on parameter sensitivity for the linear time-invariant system, S_{LC}.

Theorem 2

Suppose $A \in R^{nxn}$ is continuously differentiable with respect to

b_i at b_o. Then

$$e^{\tilde{A}_i t} = \begin{bmatrix} e^{At} & 0 \\ e^{At}_{(i)} & e^{At} \end{bmatrix} \tag{119}$$

Proof

The state transition matrix for $\tilde{A}_i$ may be uniquely determined from

$$\frac{d}{dt} e^{\tilde{A}_i t} = \tilde{A}_i e^{\tilde{A}_i t} \tag{120}$$

Partitioning $e^{\tilde{A}_i t}$ into four n x n blocks

$$e^{\tilde{A}_i t} = \begin{bmatrix} \Phi_{11}(t) & \Phi_{12}(t) \\ \Phi_{21}(t) & \Phi_{22}(t) \end{bmatrix} \tag{121}$$

it is easy to use (120) to show that

$$\frac{d}{dt} \Phi_{11}(t) = \frac{d}{dt} \Phi_{22}(t) = \frac{d}{dt} \Phi(t) = A\Phi(t) \qquad \Phi(0) = I; \tag{122}$$

$$\Phi_{12}(t) = 0 \text{ for all } t \in [0, t_f]; \tag{123}$$

and

$$\frac{d}{dt} \Phi_{21}(t) = A\Phi_{21}(t) + A_{(i)}\Phi(t) \qquad \Phi_{21}(0) = 0. \tag{124}$$

Then from Theorem 1, $\Phi_{21}(t) = \Phi_{(i)}(t)$, and the proof is complete.

Q.E.D.

Theorem 2 establishes the basic relationship from which we obtain various representations for the partial derivatives, $e^{At}_{(i)}$. It shows the fundamental importance of the augmented matrix $\tilde{A}_i$ in the determination of the i^{th} partial derivative of e^{At}. Examining this augmented matrix we see that its characteristic polynomial is

$$\tilde{\Delta}_i(q) = \det \begin{bmatrix} qI - A & 0 \\ -A_{(i)} & qI - A \end{bmatrix}$$

$$= \Delta^2(q) = \prod_{k=1}^{\rho} (q - q_k)^{2n_k} \tag{125}$$

Thus the eigenvalues of $\tilde{A}_i$ are the same as those of A and this is true for each parameter component b_i, i = 1, 2, ...p. Of course the minimal polynomials of the $\tilde{A}_i$'s may be different for each parameter component, b_i, and so let us suppose that $\tilde{A}_i$ has a $\tilde{\mu}_i$ - order minimal polynomial

$$\tilde{\psi}_i(q) = \prod_{k=1}^{\rho} (q - q_k)^{\tilde{\mu}_{i_k}} \tag{126}$$

We shall investigate the eigenvalue multiplicities, $\tilde{\mu}_{i_k}$, k = 1, 2, ...ρ, further, but for now the above notation is used to give the following result:

Corollary 2.1

$$e^{At}_{(i)} = \sum_{k=1}^{\rho} \sum_{j=0}^{\tilde{\mu}_{i_k}-1} \tilde{V}_{i_{k,j}} t^j e^{q_k t} \tag{127}$$

where

$$\tilde{V}_{i_{k,j}} \equiv [I \quad 0]_{n \times 2n} \tilde{Z}_{i_{k,j}} \begin{bmatrix} 0 \\ I \end{bmatrix}_{2n \times n} \tag{128}$$

and $Z_{i_{k,j}}$ are the components of $\tilde{A}_i$.

Proof

This is an immediate consequence of Theorem 2.

Q.E.D.

Corollary 2.1 is the primary representation which will be used in our computational applications, and Appendix A discusses some computa-

tional techniques for determining the matrices $\tilde{V}_{i_{k,j}}$. However, for investigation of some of the geometric properties of the sensitivity operators, the next corollary will find principle application.

Corollary 2.2

$$e^{At} = \sum_{j=1}^{2n} A^{j-1} a_j(t) \tag{129}$$

and for all i = 1, 2, ...p

$$e^{At}_{(i)} = \sum_{j=1}^{2n} (A^{j-1})_{(i)} a_j(t) \tag{130}$$

where the scalar functions $a_j(\cdot)$, j = 1, 2, ...2n, are determined from

$$t^j e^{q_k t} = \frac{d^j}{dq^j} [1 \quad q \quad \ldots \quad q^{2n-1}]_{q=q_k} \begin{bmatrix} a_1(t) \\ \vdots \\ a_{2n}(t) \end{bmatrix} \tag{131}$$

$k = 1, 2, \ldots \rho$

$j = 0, 1, \ldots 2n_k - 1$

Proof

The corollary follows immediately from Theorem 2 upon using the well-known result that (e.g., (Ref 51))

$$\begin{bmatrix} A & 0 \\ A_{(i)} & A \end{bmatrix}^j = \begin{bmatrix} A^j & 0 \\ (A^j)_{(i)} & A^j \end{bmatrix} \tag{132}$$

Q.E.D.

Before considering the application of Corollary 2.2 to the defining relations for the sensitivity operators, some fundamental information

concerning the partial derivatives $e^{At}_{(i)}$ can be gained by examining the minimal polynomials of $\tilde{A}_i$.

b. The Minimal Polynomial of $\tilde{A}_i$ - General Case

The following lemma summarizes a useful way to determine the minimal polynomial of $\tilde{A}_i$. Notice that the only assumption on the A matrix for this lemma is that the partial derivative $A_{(i)}$ exist. Indeed, one could substitute the matrix $A_{(i)}$ with any n x n matrix and the conclusions of the lemma would still remain valid.

Lemma 6

Let $\Delta(q)$ and $\psi(q)$ be the characteristic and minimal polynomials, respectively, of A as defined by equations (111) and (112).

i) Suppose that $g(q)$ is the polynomial such that $\Delta(q) = \psi(q)g(q)$, and define the reduced adjunct matrix[1]

$$\overline{\text{adj}}\,[qI - A] \equiv \text{adj}\,[qI - A]/g(q) \tag{133}$$

Let $g_i(q)$ be the largest common factor of both $\overline{\text{adj}}[qI - A]A_{(i)}\overline{\text{adj}}[qI - A]$ and the minimal polynomial $\psi(q)$. Then $\tilde{\psi}_i(q) = \psi^2(q)/g_i(q)$ is the minimal polynomial of $\tilde{A}_i$, equation (118).

ii) The order of $\tilde{\psi}_i(q)$ is between the bounds

$$\mu \leq \tilde{\mu}_i \leq 2\mu \tag{134}$$

and each eigenvalue multiplicity has similar bounds

$$\mu_k \leq \tilde{\mu}_{i_k} \leq 2\mu_k . \tag{135}$$

---

[1]Following Nering (Ref 21: 95) we term the "adjunct matrix" of A as the matrix adj (A) = $(A_{ij})^T$ where A_{ij} is the cofactor of the element a_{ij} of A. The adjunct matrix is also commonly referred to as the "adjoint" matrix, and we use this terminology to avoid confusion with the adjoint operator.

Proof

Part i) is derived by considering the 2n x 2n partitioned adjunct matrix

$$\text{adj} \begin{bmatrix} qI - A & 0 \\ -A_{(i)} & qI - A \end{bmatrix} = \begin{bmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{bmatrix} \tag{136}$$

Using the fact that for $q \neq q_k$ (Ref 21: 96)

$$[qI - A] \text{ adj } [qI - A] = \Delta(q)I \tag{137}$$

one may form a set of four linear equations from

$$\begin{bmatrix} qI - A & 0 \\ -A_{(i)} & qI - A \end{bmatrix} \begin{bmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{bmatrix} = \Delta^2(q) \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{138}$$

and show that

$$\text{adj} \begin{bmatrix} qI - A & 0 \\ -A_{(i)} & qI - A \end{bmatrix} =$$

$$\begin{bmatrix} \Delta(q) \text{ adj } [qI - A] & 0 \\ \text{adj } [qI - A] A_{(i)} \text{ adj } [qI - A] & \Delta(q) \text{ adj } [qI - A] \end{bmatrix} \tag{139}$$

Now g(q) is the largest common factor of every term of adj [qI - A] (Ref 21: 101), and so the largest common factor of adj $[qI - A_i]$ will be $g^2(q)g_i(q)$ where $g_i(q)$ is as described in the lemma. Then the minimal polynomial of $\tilde{A}_i$ is

$$\tilde{\psi}_i(q) = \Delta^2(q)/g^2(q)g_i(q) = \psi^2(q)/g_i(q) \tag{140}$$

Part ii) of the lemma then follows immediately from equation (140) and the definition of $g_i(q)$.

Q.E.D.

The minimal polynomial of the augmented matrix $\tilde{A}_i$ has been considered previously by Guardabassi, Locatelli, and Rinaldi (Ref 51). By examining the Jordan canonical form of $\tilde{A}_i$ they obtain similar bounds on the order and multiplicities of the minimal polynomial[1]. They obtain slightly tighter bounds on the multiplicities of the eigenvalues (they conclude that $\tilde{\mu}_{k_i} \leq 2\mu_k - 1 + \delta_{1,\mu_k}$ where μ_k is the eigenvalue multiplicity of q_k and δ_{1,μ_k} is the Kronecker delta); however, the results of Lemma 6 are completely general requiring no assumptions concerning the parameters in A other than the existence of the partial derivatives, whereas Guardabassi, et al, place the restriction on the structure of A that the dimension of its Jordan blocks and number of distinct eigenvalues are invariant in a neighborhood of the nominal parameter value, b_o. The following simple example illustrates how Lemma 6 might be utilized and shows that the restrictions assumed by Guardabassi, et al, are indeed necessary in order to obtain their slightly tighter bounds. Also, this example will provide insight into our next subsection in which we consider a special case similar to the one which Guardabassi, et al, considered.

Example

Suppose that A(b) is the 4 x 4 matrix

$$A(b) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ b_1 & 1 & 0 & 0 \\ 0 & 0 & 1 & b_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{141}$$

[1]Actually reference 51 considers the minimal polynomial of $\bar{A}$, equation (6); however, Lemma 6 is easily extended to this case and part ii) of the lemma remains true regardless.

where the nominal values of b_1 and b_2 are zero. Then $\Delta(q) = (q - 1)^4$ and $\psi(q) = (q - 1)^2$. The adjunct matrix of $[qI - A]$ is readily computed to be

$$\text{adj } [qI - A] = (q - 1)^2 \begin{bmatrix} (q - 1) & 1 & 0 & 0 \\ 0 & (q - 1) & 0 & 0 \\ 0 & 0 & (q - 1) & 0 \\ 0 & 0 & 0 & (q - 1) \end{bmatrix} \quad (142)$$

The partial derivatives of A are

$$A_{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A_{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (143)$$

and so

$$\overline{\text{adj}}\ [qI - A]A_{(1)}\overline{\text{adj}}\ [qI - A] = \begin{bmatrix} (q - 1) & 1 & 0 & 0 \\ (q - 1)^2 & (q - 1) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (144)$$

$$\overline{\text{adj}}\ [qI - A]A_{(2)}\overline{\text{adj}}\ [qI - A] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (q - 1)^2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (145)$$

Therefore, by Lemma 6 the minimal polynomial of $\tilde{A}_1$ is

$$\tilde{\psi}_1(q) = \psi^2(q)/1 = (q - 1)^4 \quad (146)$$

and of A_2 is

$$\tilde{\psi}_2(q) = \psi^2(q)/(q-1)^2 = (q-1)^2 = \psi(q). \tag{147}$$

Thus, this example has the two extremes in which $\tilde{\psi}_1(q) = \psi^2(q)$ and $\tilde{\psi}_2(q) = \psi(q)$. Notice that any value of b_1 other than zero produces the characteristic polynomial

$$\Delta_{b_1}(q) = (q-1)^2(q^2 - 2q + 1 - b_1) \tag{148}$$

and minimal polynomial

$$\psi_{b_1}(q) = (q-1)(q^2 - 2q + 1 - b_1) \tag{149}$$

whereas a perturbation of b_2 has no effect on the minimal polynomial of A.

c. Special Case - Invariance of the Minimal Polynomial of A

Assume that A has the property that the structure of its minimal polynomial is invariant for all parameter component perturbations in a neighborhood of the nominal b_o. There is little loss in generality in this case from that of the previous section, as it happens very seldom in practice that a local parameter perturbation will change the structure of the minimal polynomial. Also, this case is slightly more general than the case considered by Guardabassi, et al, (Ref 51), as the structure of the Jordan blocks of a matrix can change without affecting the minimal polynomial but not vice versa. The following theorem summaries a fundamental relationship between the minimal polynomial of $\tilde{A}_i$ and the b_i-eigenvalue sensitivities for this special case:

Theorem 3

Suppose that A has μ^{th}-order minimal polynomial, $\psi(q)$, equation (112); A is continuously differentiable with respect to b_i at b_o; and for all

$|\Delta b_i| \leq \varepsilon_i$, $A_{\Delta_i} = A(b_o + \Delta b_i e_i)$ has minimal polynomial

$$\psi_{\Delta_i}(q) = \prod_{k=1}^{\rho} (q - q_k^{\Delta_i})^{\mu_k} \tag{150}$$

where $q_k^{\Delta_i} \to q_k$ as $|\Delta b_i| \to 0$, $k = 1, 2, \ldots \rho$.

Then:

i) $$e^{At}_{(i)} = \sum_{k=1}^{\rho} (Z_{k,0})_{(i)} e^{q_k t}$$

$$+ \sum_{k=1}^{\rho} \sum_{j=1}^{\mu_k - 1} \{(Z_{k,j})_{(i)} + Z_{k,j-1}(q_k)_{(i)}\} t^j e^{q_k t}$$

$$+ \sum_{k=1}^{\rho} Z_{k,\mu_k - 1}(q_k)_{(i)} t^{\mu_k} e^{q_k t} \tag{151}$$

where $Z_{k,j}$ are the components of A and $(Z_{k,j})_{(i)}$ and $(q_k)_{(i)}$ are the component and eigenvalue sensitivities (partial derivatives) with respect to b_i. These partial derivatives exist and are continuous.

ii) The minimal polynomial of the augmented matrix $\tilde{A}_i$ is

$$\tilde{\psi}_i(q) = \prod_{k=1}^{\rho} (q - q_k)^{\tilde{\mu}_{i_k}} \tag{152}$$

where

$$\tilde{\mu}_{i_k} = \begin{cases} \mu_k \text{ if } (q_k)_{(i)} = 0 \\ \mu_k + 1 \text{ if } (q_k)_{(i)} \neq 0 \end{cases} \tag{153}$$

iii) The order of the minimal polynomial of $\tilde{A}_i$ is $\mu + \beta_i$ where β_i is the number of eigenvalues of A which have a non-zero b_i eigenvalue sensitivity.

Proof

From the component representation of e^{At} we have

$$e^{At} = \sum_{k=1}^{\rho} \sum_{j=0}^{\mu_k - 1} Z_{k,j} e^{q_k t} t^j. \tag{154}$$

Similarly, by the assumption on A_{Δ_i} it is true that

$$e^{A_{\Delta_i} t} = \sum_{k=1}^{\rho} \sum_{j=0}^{\mu_k - 1} Z_{k,j}^{\Delta_i} e^{q_k^{\Delta_i} t} t^j \tag{155}$$

where $Z_{k,1}^{\Delta_i}$ are the component matrices of A_{Δ_i}. Then by the definition of the partial derivative, we obtain:

$$e^{At}_{(i)} = \sum_{k=1}^{\rho} \sum_{j=0}^{\mu_k - 1} \lim_{|\Delta b_i| \to 0} \frac{Z_{k,j}^{\Delta_i} e^{q_k^{\Delta_i} t} - Z_{k,j} e^{q_k t}}{\Delta b_i} t_j \tag{156}$$

Since A is continuously differentiable with respect to b_i, by Theorem 1 it is true that $e^{At}_{(i)}$ exists and is continuous. Thus the limits on the right hand side of equation (156) must also exist and be continuous. By proper manipulation expression (151) readily follows. Then since the components $Z_{k,\mu_k - 1}$ cannot be zero by definition of the minimal polynomial, parts ii) and iii) follow immediately from i).

Q.E.D.

The conclusions of Theorem 3 are an extension to those of Guardabassi, et al (Ref 51) (Ref 52). In (Ref 51) upper bounds on the multiplicities of the eigenvalues of the augmented matrix $\bar{A}$, equation (7), are obtained (see the comment following Lemma 6). Then in (Ref 52) the special case in which there is no repeated eigenvalues in the A matrix is analyzed in detail. The detailed structure of the Jordan canonical form of $\bar{A}$ is given, and the structure of these blocks is dependent upon the eigenvalue sensitivities. In both papers the conclusions are reached through an

analysis of the Jordan canonical form of $\bar{A}$. Thus, Theorem 3 extends the conclusions of these previous papers by obtaining definitive relationships between the eigenvalue sensitivities and the minimal polynomial of the augmented matrix $\tilde{A}_i$ (the extension to the minimal polynomial of $\bar{A}$ is straight-forward - see Corollary 3.1), and this is accomplished for the more general case (repeated eigenvalues). Also, it is significant to note that the approach taken in deriving Theorem 3 is considerably different from that of References 51 and 52, in that a direct differentiation of e^{At} is utilized rather than the Jordan canonical form analysis used by Guardabassi, et al.

Comparing equation (127) of Corollary 2.1 with equation (151) of Theorem 3 we see that

$$\tilde{V}_{i_{k,\mu_k}} = Z_{k,\mu_k-1}(q_k)_{(i)}. \tag{157}$$

Since the n x n matrices $\tilde{V}_{i_{k,\mu_k}}$ and Z_{k,μ_k-1} may be determined independently from the components of the matrices $\tilde{A}_i$ and A, respectively, the relation (165) provides a very general and straight-forward means for determining the eigenvalue sensitivity $(q_k)_{(i)}$, itself. Most of the existing literature on eigenvalue sensitivity has concentrated on the special case in which A has distinct eigenvalues or the parameter appears linearly in one or more elements of the A matrix (Ref 43) (Ref 44). The relation (157) places only the restrictions assumed by Theorem 3. We note that Guardabassi, et al, also derive a general expression for computing the eigenvalue sensitivities based upon the Jordan blocks of the canonical form of $\bar{A}$.

Next, if the conditions of Theorem 3 are satisfied for all b_i, i = 1, 2, ...p, and if we let $\beta \leq \rho$ be the number of eigenvalues

of A which have a non-zero sensitivity for at least one b_i, i = 1, 2, ...p; then for all i = 1, 2, ...p it is possible to represent $e^{At}_{(i)}$ by a $(\mu + \beta)$ order polynomial as given by the next corollary:

Corollary 2.3

Assume the conditions of Theorem 3 are true for all parameter components, b_i, i = 1, 2, ...p. Let β be the number of eigenvalues which have at least one non-zero parameter sensitivity for b_i, i = 1, 2, ...p. Then:

i) The minimal polynomial of the n(p + 1) x n(p + 1) augmented matrix $\bar{A}$ (defined in equation (7)) is

$$\bar{\psi}(q) = \prod_{k=1}^{\rho} (q - q_k)^{\bar{\mu}_k} \tag{158}$$

where

$$\bar{\mu}_k = \begin{cases} \mu_k & \text{if } (q_k)_{(i)} = 0 \text{ for all } i = 1, 2, \ldots p \\ \mu_k + 1 & \text{if } (q_k)_{(i)} \neq 0 \text{ for at least one } i = 1, 2, \ldots p \end{cases}$$

The order of $\bar{\psi}(q)$ is $\mu + \beta \equiv \bar{\mu}$.

ii) For all i = 1, 2, ...p

$$e^{At}_{(i)} = \sum_{j=1}^{\bar{\mu}} (A^{j-1})_{(i)} \bar{\alpha}_j(t) \tag{159}$$

where the real scalar functions $\bar{\alpha}_j(\cdot)$, j = 1, 2, ...$\bar{\mu}$, are determined from

$$t^j e^{q_k t} = \frac{d^j}{dq^j} [1 \quad q \quad \ldots \quad q^{\bar{\mu}-1}]_{q=q_k} \begin{bmatrix} \bar{\alpha}_1(t) \\ \vdots \\ \bar{\alpha}_{\bar{\mu}}(t) \end{bmatrix} \tag{160}$$

k = 1, 2. ...ρ

j = 0, 1, ...$\bar{\mu}_k - 1$

Proof

Extending Theorem 2 it can readily be shown that

$$e^{\bar{A}t} = \begin{bmatrix} e^{At} & 0 & \cdots & 0 \\ e^{At}_{(1)} & e^{At} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ e^{At}_{(p)} & 0 & \cdots & e^{At} \end{bmatrix}_{n(p+1)\times n(p+1)} \tag{161}$$

If $\bar{A}$ has the minimal polynomial

$$\bar{\psi}(q) = \prod_{k=1}^{\rho} (q - q_k)^{\bar{\mu}_k} \tag{162}$$

then $e^{\bar{A}t}$ has the unique component representation

$$e^{\bar{A}t} = \sum_{k=1}^{\rho} \sum_{j=0}^{\bar{\mu}_k - 1} t^j e^{q_k t} \bar{Z}_{k,j}, \tag{163}$$

where the component matrices $\bar{Z}_{k,j}$ are partitioned so that the relationship (161) is satisfied. But by this relationship and the assumptions of the corollary it must be true that for any particular eigenvalue, q_k, $k = 1, 2, \ldots\rho$, the highest power of t in the expression (163) will be either $\mu_k - 1$ if $(q_k)_{(i)} = 0$ for all $i = 1, 2, \ldots p$ or it will be μ_k otherwise. Thus the conclusions of part i) readily follow.

Part ii) is then merely an application of the conclusions of i).

Q.E.D.

Thus Corollary 2.3 may be used in lieu of Corollary 2.2, and the number of linear equations which need to be solved (to determine the scalar functions $a_j(\cdot)$ in the former case and the functions $\bar{\alpha}_j(\cdot)$ in the latter) will be reduced from 2n to $\bar{\mu}$. If there are many repeated eigen-

values, then the difference could be significant, particularly since solving the set of required linear equations will generally involve the inversion of a generalized Vandermonde matrix of the corresponding dimension. However, the use of Corollary 2.3 requires considerably more structural knowledge than Corollary 2.2 and this structural knowledge probably far outweighs the benefits of the reduced order Vandermonde matrix.

We conclude this subsection with a simple example to help motivate and clarify the concepts which have been discussed.

Example

Consider the 2 x 2 matrix

$$A(b) = \begin{bmatrix} b_1 & b_2 \\ \sin b_3 & b_3 \end{bmatrix} \quad A(b_o) = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \tag{164}$$

and so the nominal eigenvalues of A are $q_1 = 0$ and $q_2 = 1$. Applying Corollary 2.2 we have the set of linear equations in Vandermonde matrix form

$$\begin{bmatrix} 1 \\ t \\ e^t \\ te^t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} a_1(t) \\ a_2(t) \\ a_3(t) \\ a_4(t) \end{bmatrix} \tag{165}$$

Then it is fairly easy to show

$$a_1(t) = 1 \quad a_2(t) = t \quad a_3(t) = -3 - 2t + 3e^t - te^t \tag{166}$$

$$a_4(t) = 2 + t - 2e^t + te^t \tag{167}$$

Then using Corollary 2.2, combining like functions of t, one can represent

the partial derivatives in the form of Corollary 2.1 as follows:

$$e^{At} = Z_{10} + Z_{20}\, e^{t} = \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} e^{t} \tag{168}$$

$$e^{At}_{(1)} = \tilde{V}_{1_{1,0}} + \tilde{V}_{1_{1,1}} t + \tilde{V}_{1_{2,0}} e^{t} + \tilde{V}_{1_{2,1}} te^{t}$$

$$= \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} e^{t} + \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} te^{t} \tag{169}$$

$$e^{At}_{(2)} = \tilde{V}_{2_{1,0}} + \tilde{V}_{2_{1,1}} t + \tilde{V}_{2_{2,0}} e^{t} + \tilde{V}_{2_{2,1}} te^{t}$$

$$= \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} e^{t} \tag{170}$$

$$e^{At}_{(3)} = \tilde{V}_{3_{1,0}} + \tilde{V}_{3_{1,1}} t + \tilde{V}_{3_{2,0}} e^{t} + \tilde{V}_{3_{2,1}} te^{t}$$

$$= \begin{bmatrix} 2 & 6 \\ -1 & -2 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & -1 \end{bmatrix} t + \begin{bmatrix} -2 & -6 \\ 1 & 2 \end{bmatrix} e^{t} + \begin{bmatrix} 2 & 4 \\ 0 & 0 \end{bmatrix} te^{t} \tag{171}$$

Then using equation (157) to give the expressions that

$$\tilde{V}_{i_{1,1}} = (q_1)_{(i)}\, Z_{1,0} \tag{172}$$

$$\tilde{V}_{i_{2,1}} = (q_2)_{(i)}\, Z_{2,0} \tag{173}$$

we see that the eigenvalue sensitivities are given by

$$(q_1)_{(1)} = 0 \qquad (q_1)_{(2)} = 0 \qquad (q_1)_{(3)} = -1 \tag{174}$$

$$(q_2)_{(1)} = 1 \qquad (q_2)_{(2)} = 0 \qquad (q_2)_{(3)} = 2 \tag{175}$$

Then since both q_1 and q_2 have a non-zero sensitivity for at least one parameter component, b_1, b_2, or b_3, the minimal polynomial of $\bar{A}$ is given by

$$\bar{\psi}(q) = (q - 0)^{1+1}\,(q - 1)^{1+1} = q^2(q - 1)^2. \tag{176}$$

Thus the order of the minimal polynomial of $\bar{A}$ is $\bar{\mu} = \mu + 2 = 4$.

d. Parameter Sensitivity Operators for System S_{LC}

In the previous sections a number of basic mathematical representations of the partial derivatives $e^{At}_{(i)}$ were developed. These will now be used to provide a convenient operator description of the sensitivities for the system S_{LC}. In particular, by directly utilizing equations (96) (100) and Corollary 2.2 we obtain the fundamental algebraic description of these parameter sensitivity operators[1]:

Theorem 4

For the system S_{LC} at the a priori parameter value $b_o \in R^p$ it is true that

i) $$y_{z.i.}(t) = \sum_{j=1}^{2n} CA^{j-1} da_j(t) \tag{177}$$

$$y_{z.s.}(t;u) = \sum_{j=1}^{2n} CA^{j-1}B \int_0^t a_j(t - s)u(s)ds \tag{178}$$

$$y(t;u) = y_{z.i.}(t) + y_{z.s.}(t;u) \tag{179}$$

---

[1]Note that Corollary 2.4 or any of the other equivalent polynomial representations of $e^{At}_{(i)}$ could likewise be utilized here.

ii) $$v^{(i)}_{z.i.}(t) = \sum_{j=1}^{2n} (CA^{j-1}d)_{(i)} a_j(t) \tag{180}$$

$$v^{(i)}_{z.s.}(t;u) = \sum_{j=1}^{2n} (CA^{j-1}B)_{(i)} \cdot \int_0^t a_j(t-s)u(s)ds \tag{181}$$

$$v^{(i)}(t;u) = v^{(i)}_{z.i.}(t) + v^{(i)}_{z.s.}(t;u) \tag{182}$$

$i = 1, 2, \ldots p$

where the scalar functions, $a_j(t)$, $j = 1, 2, \ldots 2n$, are uniquely determined by the 2n linear equations (131).

Proof

The theorem is a direct application of Corollary 2.2 to the defining equations (96) - (110).

Q.E.D.

Theorem 4 provides a useful representation of the parameter sensitivity operators for the linear time-invariant system S_{LC}. The geometric aspects of this description are examined in the next section in conjunction with the systems properties of insensitivity, sensitivity controllability, and identifiability. Theorem 4 is also a convenient form for digital computer applications, and some of the computational aspects of this description are examined in Appendix A. Note that the entire set of first order sensitivities may be obtained with at most 2nr ordinary integrals (Appendix A describes how the convolutions indicated in equations (178) and (181) may be transformed into time varying combinations of quadrature integrals). The computational utility of this representation will be seen even more clearly when it is used to form the basis of algorithms for quasilinearization

parameter identification (Section V), minimum sensitivity control design (Section VI), and optimal sensitivity design for improved identification (Section VII). Finally, we note that a similar polynomial description for the second and higher order sensitivities may easily be obtained, and a discussion of this is presented in Appendix B.

3. SUMMARY

In this section we have used an operator description of the system output to obtain the sensitivity operators for linear ordinary differential equations systems - both time-varying and time-invariant ones. These results illustrate the operator approach used in Section II for the Hilbert space system S_H. The key to making the operator approach practical is obtaining the partial derivatives of the state transition matrix with respect to the parameter components. Expressions for these partial derivatives are derived for the time-varying system S_{TV} (Theorem 1); however, for merely computing the output sensitivity operators their use is probably less desirable than computing the sensitivity operators from the "sensitivity system" differential equations.

For the time-invariant system S_{LC} this is not at all the case. Here polynomial expressions for the partial derivatives of e^{At} are derived enabling both geometric insights and practical computational techniques to be obtained. The complete system output and all p output parameter sensitivity operators can thus be obtained through matrix manipulations and the solution of 2nr quadrature integrals where n is the state dimension and r is the control dimension. If mild structural restrictions on the minimal polynomial of A are assumed in a neighborhood of the nominal parameter vector, then explicit relationships between the modes of the sensitivity operators and the eigenvalue sensitivities are obtained. These expressions also result in a new

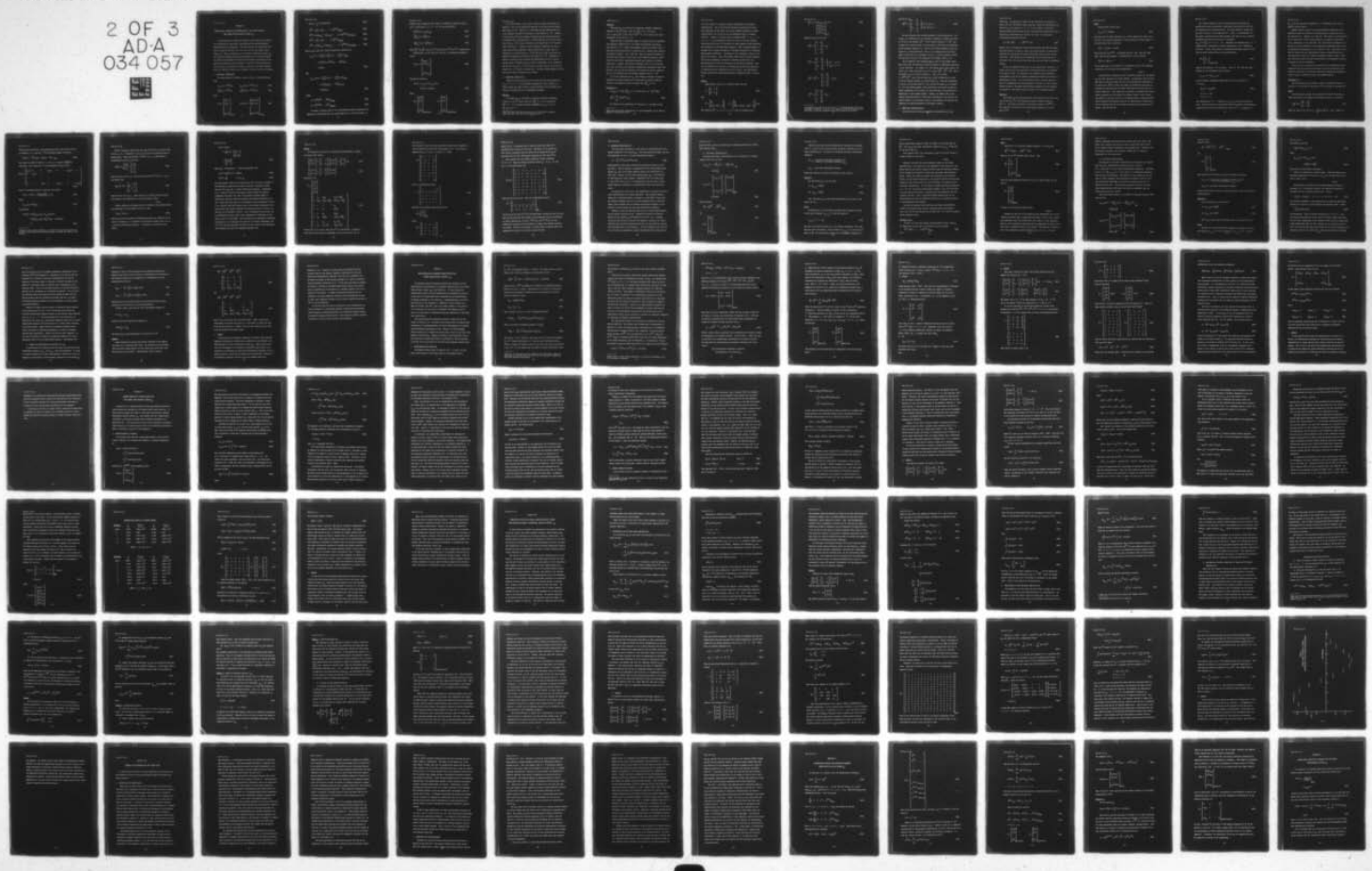
general method for computing the eigenvalue sensitivities themselves. These structural properties of the sensitivity operators are extensions of those obtained previously through Jordan canonical form analysis (Ref 51) (Ref 52).

In the next four sections we use the approach of Section II to examine the sensitivity-related system properties of insensitivity, sensitivity controllability, and identifiability (Section IV) and to develop computational algorithms for quasilinearization (Section V), mininum sensitivity control design (Section VI), and optimal design of the sensitivity operators to enhance parameter identification (Section VII). We concentrate on the linear time-invariant system and the application of the algebraic representation obtained in Section III.2.

AD-A034 057 AIR FORCE AVIONICS LAB WRIGHT-PATTERSON AFB OHIO F/G 12/1
SENSITIVITY OPERATORS AND ASSOCIATED SYSTEM CONCEPTS FOR LINEAR--ETC(U)
JUL 76 J G REID

UNCLASSIFIED AFAL-TR-76-118 NL

2 OF 3
AD-A
034 057



Section IV

INSENSITIVITY, SENSITIVITY CONTROLLABILITY, AND IDENTIFIABILITY FOR LINEAR TIME-INVARIANT SYSTEMS, S_{LC}

In this section the sensitivity-related system properties of parameter insensitivity, sensitivity controllability and system parameter identifiability are considered. The basic concepts for these properties were discussed in Sections II.3, II.4, and II.7 for the Hilbert space system, S_H. Here we concentrate on the constant coefficient system, S_{LC}. Our reason for grouping these three properties in this section is based upon their direct geometric reliance on the algebraic description of the sensitivity operators given by Theorem 4. Also, all three of these system properties are intimately related to each other, and these relationships are made more evident by the algebraic conditions which are obtained.

1. NOTATIONAL CONVENTIONS

The representation of Theorem 4 may be written in the matrix-vector form:

$$y_{z.i.}(t) = E^{(0)}a(t) \qquad y_{z.s.}(t;u) = G^{(0)}g(t;u) \tag{183}$$

$$v^{(i)}_{z.i.}(t) = E^{(i)}a(t) \qquad v^{(i)}_{z.s.}(t;u) = G^{(i)}g(t;u) \tag{184}$$

where

$$a(t) = \begin{bmatrix} a_1(t) \\ a_2(t) \\ \vdots \\ a_{2n}(t) \end{bmatrix}_{2nx1} \qquad a(t)*u(t) = \begin{bmatrix} a_1(t)u(t) \\ a_2(t)u(t) \\ \vdots \\ a_{2n}(t)u(t) \end{bmatrix}_{2nrx1} \tag{185}$$

$$g(t;u) = \int_0^t a(t-s)*u(s)ds \tag{186}$$

$$E^{(0)} = [Cd \mid CAd \mid \ldots \mid CA^{2n-1}d]_{mx2n} \tag{187}$$

$$E^{(i)} = [(Cd)_{(i)} \mid (CAd)_{(i)} \mid \ldots \mid (CA^{2n-1}d)_{(i)}]_{mx2n} \tag{188}$$

$$G^{(0)} = [CB \mid CAB \mid \ldots \mid CA^{2n-1}B]_{mx2nr} \tag{189}$$

$$G^{(i)} = [(CB)_{(i)} \mid (CAB)_{(i)} \mid \ldots \mid (CA^{2n-1}B)_{(i)}]_{mx2nr} \, , \tag{190}$$

Then we may write the "total" sensitivity operators as

$$\begin{aligned} V_{z.i.}(t) &\equiv [v^{(1)}_{z.i.}(t)v^{(2)}_{z.i.}(t)\ldots v^{(p)}_{z.i.}(t)]_{mxp} \\ &= [E^{(1)}a(t)\ E^{(2)}a(t)\ldots E^{(p)}a(t)] \\ &= E*a(t) \end{aligned} \tag{191}$$

and

$$\begin{aligned} V_{z.s.}(t;u) &\equiv [v^{(1)}_{z.s.}(t;u)\ldots v^{(p)}_{z.s.}(t;u)]_{mxp} \\ &= [G^{(1)}g(t;u)\ldots G^{(p)}g(t;u)] \\ &= G*g(t;u) \end{aligned} \tag{192}$$

where

$$E = [E^{(1)}E^{(2)}\ldots E^{(p)}]_{mx2np} \tag{193}$$

$$G = [G^{(1)}G^{(2)}\ldots G^{(p)}]_{mx2nrp} \tag{194}$$

In order to discuss some of the relationships between insensitivity, sensitivity controllability, and identifiability, it will be helpful to

consider these properties with respect to individual component outputs, $y_j(\cdot) \in L_2(0,t_f;R)$, $j = 1, 2, \ldots m$. We use the notation

$$v_j^{(i)}(t;u) = y_j(t;u)_{(i)} \tag{195}$$

$$v_{z.i._j}^{(i)}(t) = E_j^{(i)} a(t) \tag{196}$$

$$v_{z.s._j}^{(i)}(t;u) = G_j^{(i)} g(t;u) \tag{197}$$

where $E_j^{(i)}$ and $G_j^{(i)}$ are the j^{th} row vectors of $E^{(i)}$ and $G^{(i)}$, respectively.

Additionally, we will deal with the $m(p + 1)$ dimensioned augmented vector

$$Y(t;u) = \begin{bmatrix} y(t;u) \\ v^{(1)}(t;u) \\ \vdots \\ v^{(p)}(t;u) \end{bmatrix} \tag{198}$$

This may be written as

$$Y(t;u) = Y_{z.i.}(t) + Y_{z.s.}(t;u)$$

$$= \bar{E}\, a(t) + \bar{G}\, g(t;u) \tag{199}$$

where

$$\bar{E} = \begin{bmatrix} E^{(0)} \\ E^{(1)} \\ \vdots \\ E^{(p)} \end{bmatrix}_{m(p+1)x2n}, \qquad \bar{G} = \begin{bmatrix} G^{(0)} \\ G^{(1)} \\ \vdots \\ G^{(p)} \end{bmatrix}_{m(p+1)x2nr} \tag{200}$$

As a final comment, notice that we have not shown the explicit dependency of any of the sensitivity operators or matrices upon the parameter vector $b \in R^p$. Naturally, if these matrices and sensitivities are to be computed, they must be evaluated at some nominal $b_o \in R^p$. However, in this section one may be equally concerned with the system properties of insensitivity, controllability, or identifiability for just one particular value $b_o \in R^p$ (a local condition) or for nearly all $b \in R^p$ (a "structural" condition[1]). Because we obtain algebraic conditions for each of these three system properties, these conditions may be made local by evaluation at one particular $b_o \in R^p$ or they may be made "structural" by leaving the appropriate matrices as functions of $b \in R^p$. Therefore, we adopt the convention that theorems and definitions will all be stated without regard to whether they are local or structural; if the conditions are met for just one $b_o \in R^p$ the property is local, but if it is true for almost all $b \in R^p$ then it is structural. The later simple examples of this section will better illustrate this concept.

2. PARAMETER INSENSITIVITY

Using the previous notation and the fact that the scalar functions $a_j(\cdot)$, $j = 1, 2, \ldots 2n$ are linearly independent on every interval of finite length (and hence non-zero almost everywhere) the following conditions for insensitivity can be immediately obtained.

Theorem 5

The system S_{LC} is zero-input b_i-component parameter insensitive relative to the j^{th} output (that is, $v^{(i)}_{z.i._j}(\cdot) = 0$) if and only if $E^{(i)}_j = 0$. It is totally zero-input insensitive ($V_{z.i.}(\cdot) = 0$) if and only if $E = 0$.

---

[1]Gupta and Hall (Ref 108) introduce such a "structural" concept in their discussion of sensitivity controllability.

Theorem 6

The system S_{LC} is zero-state b_i-component parameter insensitive relative to the j^{th} output (that is, $v^{(i)}_{z.s._j}(\cdot;u) = 0$ for all $u(\cdot) \in L_2(0,t_f;R^r)$) if and only if $G^{(i)}_j = 0$. It is totally zero-state insensitive if and only if $G = 0$.

There has been a significant amount of research into the conditions for parameter insensitivity in the linear time-invariant system S_{LC}. (See Section I.1.a.) Indeed, the essential results of Theorem 5 and 6 have been derived previously by "canonical form - minimal order sensitivity system" analysis. (e.g., (Ref 61)) They are stated here mainly to illustrate how Theorem 4 can be used to give immediate geometric conditions for which fairly complicated analysis were required in previous derivations. However, the results of Theorems 5 and 6 are new in that they treat zero-input and zero-state insensitivity separately rather than jointly[1], and they add the generality of allowing the initial condition vector to be dependent upon the unknown parameter vector, $b \in R^p$.

If d is not dependent upon b_i, then Theorem 5 provides the means to determine the largest subspace of R^n from which an initial condition vector may be selected in order to ensure that $v^{(i)}_{z.i._j}(\cdot) = 0$;

Corollary 5.1

If $d_{(i)} = 0$, then $v^{(i)}_{z.i._j}(\cdot) = 0$ if and only if $d \in Z^{(i)}_j$ where

$$Z^{(i)}_j = \bigcap_{k=1}^{2n} N\{(C_j A^{k-1})_{(i)}\} \tag{201}$$

To illustrate the application of Corollary 5.1, one might utilize

---

[1]Note that "zero-state insensitivity" is what Guardabassi, et al (Ref 61) define as "hypo-insensitivity"

it in the design of "transient response" experiments for parameter identification. One of the problems frequently encountered in parameter estimation is that there are too many unknown parameters to effectively estimate all of them in any one experiment (see, for example, (Ref 85)). However, by constructing the subspaces $Z_j^{(i)}$, i = 1, 2, ...p, j = 1, 2, ...m, the possibility exists to select an initial condition vector which will zero the "influence" of some parameters while maximizing the "influence" of the remaining parameters. Such a procedure might be a possible way to reduce the number of "exciting" parameters in a parameter identification experiment. An example of this technique is given at the end of this section. Also, this example shows that the zeroing subspaces $Z_j^{(i)}$ provide fundamental information concerning insensitivity and identifiability for both the zero-input and zero-state responses. This is an illustration of just one of the many ways in which the properties of insensitivity and identification can be related. The relationship between insensitivity and identifiability is examined further in Section IV.4.

Example

Suppose that C is the 2 x 2 identity matrix and that

$$A = \begin{bmatrix} b_1 & 0 \\ b_2 & b_3 \end{bmatrix} \tag{202}$$

Then

$$A^2 = \begin{bmatrix} b_1^2 & 0 \\ b_1b_2 + b_2b_3 & b_3^2 \end{bmatrix} \quad A^3 = \begin{bmatrix} b_1^3 & 0 \\ b_1^2b_2 + b_1b_2b_3 + b_2b_3^2 & b_3^3 \end{bmatrix} \tag{203}$$

and subspaces $Z_j^{(i)}$, j = 1, 2, i = 1, 2, 3, may be determined from

$$Z_j^{(i)} = N \begin{bmatrix} (C_jI)_{(i)} \\ (C_jA)_{(i)} \\ (C_jA^2)_{(i)} \\ (C_jA^3)_{(i)} \end{bmatrix}_{4x2} \subset R^2 \tag{204}$$

Computing these we see that[1]

$$Z_1^{(1)} = N \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ . & 0 \\ . & 0 \end{bmatrix} = \{e_2\} \tag{205}$$

$$Z_2^{(1)} = N \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ b_2 & 0 \\ . & 0 \end{bmatrix} = \begin{cases} \{e_2\} & b_2 \neq 0 \\ \{e_1e_2\} & b_2 = 0 \end{cases} \tag{206}$$

$$Z_1^{(2)} = N_1^{(3)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \{e_1e_2\} \tag{207}$$

$$Z_2^{(2)} = N \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ . & 0 \\ . & 0 \end{bmatrix} = \{e_2\} \tag{208}$$

---

[1]To simplify the notation we will use a " " in places in which there is redundant information. Also e_1 and e_2 denote the Euclidean basis vectors of R^2 and $\{.\}$ denotes the span of the vectors within the brackets.

$$Z_2^{(3)} = N \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ b_2 & . \\ . & . \end{bmatrix} = \begin{cases} \{0\} \text{ if } b_2 \neq 0 \\ \{e_1\} \text{ if } b_2 = 0 \end{cases} \tag{209}$$

We may interpret these zeroing subspaces in the following way: the output y_j is insensitive to the parameter component b_i whenever the state $x \in R^2$ is in the subspace $Z_j^{(i)}$. If the initial condition $d \in Z_j^{(i)}$, and if there is no control input, then the output y_j will never be influenced by the parameter component b_i. Notice that we have determined these subspaces on a "structural" basis; that is, the subspaces have been determined as functions of the parameter values, b_i, and the subspaces change for particular values of the parameters (for example, when $b_2 = 0$).

As we commented, these subspaces may be used to help design experiments which will systematically reduce the number of influencing components on each experiment. For example, for this simple second order system, the output y_1 is never influenced by b_2 or b_3 (since $Z_1^{(2)} = Z_1^{(3)} = R^2$), and so the output y_1 can be used to isolate b_1. Also, if we select $d \in Z_2^{(1)} \cap Z_2^{(2)} = \{e_2\}$ then the output y_2 is only influenced by b_3, and so we may uniquely estimate b_3 from the output y_2. Then once estimating b_1 and b_3, we may then select $d \in \{e_1\}$ and use the output y_2 to estimate b_2. For this simple example, these conclusions could easily be made by direct examination of the system, but this illustrates the procedure for application to more complex systems. Finally, we comment that a more complex system may prevent a "structural" examination of the various subspaces; however, for particular values of the parameters the matrices and subspaces are easily determined on the digital computer.

3. SENSITIVITY CONTROLLABILITY

In this section we discuss sensitivity controllability for the

system S_{LC}. The approach is based upon the definitions and results of Section II.4 for the Hilbert space system S_H. Since the "sensitivity system" for S_{LC} (see equation (6) - (7)) is a linear time-invariant ordinary differential equation system, controllability properties of this system may be determined from the $n(p + 1) \times n(p + 1)r$ dimensioned "controllability matrix" (Ref 33)

$$[\bar{B} \mid \bar{A}\bar{B} \mid \bar{A}^2\bar{B} \mid \ldots \mid \bar{A}^{n(p+1)-1}\bar{B}]. \tag{210}$$

Indeed, this is the approach taken by Gupta and Hall (Ref 108) and Gupta and Mehra (Ref 84) in order to reduce the required number of differential equations for calculating the sensitivities when there are zero initial conditions. They examine linear independence of the column vectors of the matrix (210) and conclude that there are never more than nr independent columns.

We use the algebraic description of the sensitivities, Theorem 4, to obtain a matrix test for sensitivity controllability which is equivalent to the matrix (210). For consistency, we consider output sensitivity controllability and use the operator definitions presented in Section II.4 (Definitions 4 and 5). As in the previous section, the main intention here is to demonstrate how the description of Section III can be conveniently utilized to determine controllability properties of the sensitivity operators. The central result is given in the following theorem:

Theorem 7

The system S_{LC} is output sensitivity controllable if and only if the rank of the $m(p + 1) \times 2nr$ dimensioned matrix $\bar{G}$, equation (200) is $m(p + 1)$. The output sensitivity controllable subspace of $R^{m(p+1)}$ is the range space of $\bar{G}$ whereas the uncontrollable subspace is the null space of $\bar{G}^T$.

Proof

From equation (199) we have

$$Y_{z.s.}(t,u) = \bar{G}g(t;u) \tag{211}$$

Then since the 2n scalar functions $a_j(\cdot)$, which comprise the vector function $a(\cdot)$ are linearly independent on every interval of finite length, it is easily shown that the operator

$$g(t;\cdot) = \int_0^t a(t - s)(\cdot)ds \tag{212}$$

maps $L_2(0,t;R^r)$ onto R^{2nr}. (Luenberger (Ref 20: 56)) Thus the range space (controllable subspace - see Definition 4) of the operator

$$\bar{W}(t)(\cdot) = \bar{G}g(t;\cdot) \tag{213}$$

is the range space of the matrix $\bar{G}$. By Definition 4 and 5 the remaining conclusions of the theorem follow immediately.

Q.E.D.

Controllability properties of the "sensitivity system" for the system S_{LC} are quite important for several reasons. As mentioned, Gupta and Hall (Ref 108) and Gupta and Mehra (Ref 84) use controllability properties of the matrix pair $(\bar{A},\bar{B})$ to reduce the number of differential equations for computing the sensitivities for linear time-invariant systems with zero initial conditions. Their method may be extended to systems with non-zero initial conditions by the addition of suitable "pseudo" control inputs. This method of reduction results in the minimum number of required differential equations and eliminates some of the computational problems which can be associated with the use of special canonical form reduction techniques. (Ref 45) (Ref 47)

In a similar manner, we may use the sensitivity controllability properties to reduce computational requirements associated with the operator description of the sensitivities. To be more explicit, suppose that the rank of the $m(p + 1)x2nr$ matrix $\bar{G}$ is γ. Then γ is the dimension of the output sensitivity controllable subspace, and there are but γ linearly independent functions from the $m(p + 1)$ functions $\{y_{z.s.}(\cdot), v^{(i)}_{z.s.}(\cdot),\ i = 1, 2, \ldots p\}$. Therefore, the remaining $m(p + 1) - \gamma$ functions may be determined as linear combinations of the γ independent functions. Indeed, there exists a non-unique transformation matrix K such that the matrix product $K\bar{G}$ has the partitioned form

$$K\bar{G} = \begin{bmatrix} \bar{G}_c \\ - \\ 0 \end{bmatrix}_{m(p+1)x2nr} \tag{214}$$

where $\bar{G}_c$ has dimension $\gamma x\ nr$ and rank γ. (Ref 21: 54) Then the first γ elements of the transformed vector function

$$Y_{z.s.}(\cdot;u) = KY_{z.s.}(\cdot;u) \tag{215}$$

will be linearly independent combinations of the 2nr functions of $t \in [0,t_f]$

$$\int_0^t a_j(t - s)u_k(s)ds \tag{216}$$

$j = 1, 2, \ldots 2n$

$k = 1, 2, \ldots r.$

The remaining $m(p + 1) - \gamma$ elements of $Y_{z.s.}(\cdot;u)$ are then identically zero. This fact shows the relationship between sensitivity controllability properties and zero-state insensitivity; the $m(p + 1) - \gamma$ elements of

$Y_{z.s.}(\cdot;u)$ are zero-state insensitive (i.e., identically zero) for all possible control inputs.

Another important aspect of the controllability properties of the sensitivity system is associated with making $Y(t_f;u)$ identically zero at the terminal time. (Kalman, et al (Ref 17)) This aspect of controllability of the sensitivity system has been examined by a number of researchers (e.g., Guardabassi, et al, (Ref 52)), and, in particular, sufficiency conditions for uncontrollability of the sensitivity system have been an area of considerable investigation, (e.g., (Refs 48 - 53)). Guardabassi, et al, (Ref 52) present the most comprehensive results on uncontrollability of the sensitivity system, and a general statement of their results is that the sensitivity system is always uncontrollable if the parameter dimension p exceeds the control dimension r.

We may use the controllability properties of the sensitivity system along with our matrix-operator description to give the following result concerning necessary and sufficient conditions in making the sensitivity vector, $Y(t_f;u)$, identically zero at the terminal time t_f:

Corollary 7.1

For the system S_{LC} there exists a control $u^* \varepsilon L_2(0,t_f;R^r)$ such that $Y(t_f;u^*) = 0$ if and only if the rank of $[\bar{E}\ \ \bar{G}]$ is equal to the rank of $\bar{G}$.

Proof

For sufficiency, suppose that the rank of $[\bar{E}\ \ \bar{G}]$ and $\bar{G}$ are both γ_1. Then there exists a non-singular transformation matrix, K_1, such that

$$K_1[\bar{E}\ \ \bar{G}] = \begin{bmatrix} \bar{E}_1 & \bar{G}_1 \\ 0 & 0 \end{bmatrix} \tag{217}$$

where the rank of $[\bar{E}_1\ \ \bar{G}_1]$ is γ_1, (Nering (Ref 21: 54)). Define the

2nrx2nr positive definite, partitioned matrix M_c^* in which each partition is dimension r x r, and the i, j^{th} partitioned element is given by

$$M_c^*(i,j) = \int_0^{t_f} a_i(t_f - t)a_j(t_f - t)dt \; I_{rxr} \tag{218}$$

Then since the rank $\bar{G}$ = rank $\bar{G}_1 = \gamma_1$, the γ_1 x γ_1 matrix $(\bar{G}_1 M_c^* \bar{G}_1^T)$ is invertible. Next define the r x 2nr dimensioned matrix function

$$\bar{a}^*(t) \equiv \begin{bmatrix} a_1(t) & 0 & \cdots & 0 & a_2(t) & \cdots & 0 & \cdots & a_{2n}(t) & \cdots & 0 \\ 0 & a_1(t) & & \cdot & \cdot & & & & \cdot & & \\ \cdot & & & \cdot & \cdot & & & & \cdot & & \\ \cdot & & & \cdot & \cdot & & & & \cdot & & \\ 0 & \cdots & & a_1(t) & 0 & \cdots & a_2(t) & \cdots & 0 & \cdots & a_{2n}(t) \end{bmatrix} \tag{219}$$

Then it is straight-forward to show that the control[1]

$$u^*(t) = -\bar{a}^*(t_f - t)\bar{G}_1^T(\bar{G}_1 M_c^* \bar{G}_1^T)^{-1} Y_{z.i.}(t_f) \tag{220}$$

where

$$Y_{z.i.}(t_f) = \bar{E}_1 a(t_f) \tag{221}$$

will satisfy

$$\begin{aligned} Y(t_f;u^*) &= K_1^{-1}(Y_{z.i.}(t_f) + Y_{z.s.}(t_f;u^*)) \\ &= K_1^{-1}(Y_{z.i.}(t_f) + \bar{G}_1 \int_0^{t_f} a(t_f - t)*u^*(t)dt) \\ &= 0 \end{aligned} \tag{222}$$

---

[1]Indeed, by the results of Kalman, et al (Ref 17) it may be shown that u^* has norm strictly less than any other control for which the relation (222) is true.

Now for necessity, suppose that the rank of $[\bar{E}\ \bar{G}]$ is γ_1 and the rank of $\bar{G}$ is $\gamma_2 < \gamma_1$. Furthermore, let K_1, $\bar{G}_1$, and $\bar{E}_1$ be the same matrices as defined above. Then since rank $\bar{G}_1$ = rank $\bar{G}$ = $\gamma_2 < \gamma_1$, there exists a transformation matrix K_2 such that

$$K_2 K_1 \bar{G} = K_2 \begin{bmatrix} \bar{G}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \bar{G}_2 \\ 0 \\ 0 \end{bmatrix} \tag{223}$$

where the rank of $\bar{G}_2$ is γ_2. Then since the rank of $[\bar{E}\ \bar{G}]$ is $\gamma_1 > \gamma_2$, we are assured that

$$K_2 K_1 [\bar{E}\ \bar{G}] = \left[\begin{array}{c:c} & \bar{G}_2 \\ \bar{E}_2 & \\ & 0 \\ \hdashline 0 & 0 \end{array}\right] \tag{224}$$

where the rank of $\bar{E}_2$ is γ_1. Hence there is no control which will bring the system to zero conditions in a finite interval of time.

Q.E.D.

Lastly, assume for the moment that the C matrix is the identity matrix and independent of the parameter vector b so that

$$Y(t;u) = X(t;u). \tag{225}$$

Then we may use the non-unique transformation matrix, K_1, described in the proof of Corollary 7.1 to transform the "sensitivity system" into a set of γ_1 non-zero differential equations. In particular, we define the new "state"

$$\dot{X}(t) = K_1 X(t)$$

$$= K_1 [\bar{E} \;\; \bar{G}] \begin{bmatrix} a(t) \\ g(t;u) \end{bmatrix}$$

$$= \begin{bmatrix} \dot{X}_1(t) \\ - - - \\ 0 \end{bmatrix} \tag{226}$$

where $X_1(t)$ is dimension γ_1. Then from equation (89)

$$\dot{X}(t) = K_1 A K_1^{-1} X(t) + K_1 \bar{B} u(t) \tag{227}$$

$$X(0) = K_1 \bar{d} \qquad t \in [0, t_f] \tag{228}$$

But from equation (226) we are assured that only the first γ_1 elements of the differential equation (227) will be non-zero. Therefore, we may generate X(t) with only γ_1 linear differential equations. Furthermore, if γ_1 = rank $[\bar{E} \;\; \bar{G}]$ and if the separate control inputs are linearly independent functions, then this is the absolute minimum number of differential equations which will generate the complete sensitivity system. Notice that this result is an extension of the conclusions of Gupta, et al, (Ref 84) (Ref 108) from the case of zero initial conditions to the case in which the initial conditions are arbitrary and may, in fact, be a function of the parameter vector b. Also, we emphasize the fact that the K_1 transformation matrix always exists regardless of conditions on the A matrix, and it is non-unique. One convenient way to determine a transformation matrix K_1 is to determine the first set of linearly independent rows of $[\bar{E} \;\; \bar{G}]$, and then determine the matrix K_1 which leaves these independent rows unaltered and zeros the remaining dependent rows.

Example

To illustrate the results on sensitivity controllability, consider the second order system:

$$\begin{bmatrix} \dot{x}_1(t;u) \\ \dot{x}_2(t;u) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \end{bmatrix} + \begin{bmatrix} 0 \\ b_3 \end{bmatrix} u(t) \tag{229}$$

$$\begin{bmatrix} y_1(t;u) \\ y_2(t;u) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \end{bmatrix} \tag{230}$$

The matrix $\overline{G}$ is

$$\overline{G} = \begin{bmatrix} G^{(0)} \\ G^{(1)} \\ G^{(2)} \\ G^{(3)} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & b_3 & b_2b_3 & b_1b_3 + b_2^2b_3 \\ b_3 & b_2b_3 & b_1b_3 + b_2^2b_3 & 2b_1b_2b_3 + b_2^3b_3 \\ 0 & 0 & 0 & b_3 \\ 0 & 0 & b_3 & 2b_2b_3 \\ 0 & 0 & b_3 & 2b_2b_3 \\ 0 & b_3 & 2b_2b_3 & 2b_1b_3 + 3b_2^2b_3 \\ 0 & 1 & b_2 & b_1 + b_2^2 \\ 1 & b_2 & b_1 + b_2^2 & 2b_1b_2 + b_2^3 \end{bmatrix} \tag{231}$$

Notice that if b_3 is zero, then only $v^{(3)}$ is controllable, Otherwise, there will be four linearly independent rows of the matrix $\overline{G}$. and so

four elements of the zero-state sensitivity system can be computed as linear combinations of the four "controllable" ones. For instance, suppose that $b_1 = -2$, $b_2 = -3$, and $b_3 = 1$. Then

$$\bar{G} = \begin{bmatrix} 0 & 1 & -3 & 7 \\ 1 & -3 & 7 & -15 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 1 & -6 \\ 0 & 1 & -6 & 23 \\ 0 & 1 & -3 & 7 \\ 1 & -3 & 7 & -15 \end{bmatrix} \tag{232}$$

and the transformation matrix

$$K = \left[\begin{array}{cccc:c} & & & & 0 \\ & & I & & \\ \hdashline 0 & 0 & 0 & -1 & I \\ -1 & 0 & 2 & 3 & \\ -1 & 0 & 0 & 0 & \\ 0 & -1 & 0 & 0 & \end{array}\right] \tag{233}$$

provides the non-unique "controllable" form

$$K\bar{G} = \begin{bmatrix} \bar{G}_c \\ 0 \end{bmatrix}_{8x4} \tag{234}$$

where

$$\bar{G}_c = \begin{bmatrix} 0 & 1 & -3 & 7 \\ 1 & -3 & 7 & -15 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -6 \end{bmatrix} \tag{235}$$

Notice that K is selected here to leave the first four rows of $\bar{G}$ unaltered while zeroing the last four. Therefore, if the system has zero initial conditions then the entire set of sensitivities can be generated from the four differential equations for x(t) and $\xi^{(1)}(t)$.

Next, suppose that the example system has initial conditions $d^T = [1 \;\; 0]$. Using the nominal parameter values $b_1 = -2$, $b_2 = -3$, and $b_3 = 1$ it is easily shown that

$$[\bar{E} \;\; \bar{G}] = \begin{bmatrix} 1 & 0 & -2 & 6 & 0 & 1 & -3 & 7 \\ 0 & -2 & 6 & -14 & 1 & -3 & 7 & -15 \\ 0 & 0 & 1 & -3 & 0 & 0 & 0 & 1 \\ 0 & 1 & -3 & 5 & 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & -2 & 0 & 0 & 1 & -6 \\ 0 & 0 & -2 & 12 & 0 & 1 & -6 & 23 \\ 0 & 0 & 0 & 0 & 0 & 1 & -3 & 7 \\ 0 & 0 & 0 & 0 & 1 & -3 & 7 & -15 \end{bmatrix} \tag{236}$$

which has the rank of six. Then the transformation matrix

$$K_1 = \left[\begin{array}{cccccc:c} I & & & & & & 0 \\ \hdashline 0 & 0 & -2 & 0 & -3 & -1 & I \\ 0 & -1 & 0 & -2 & 2 & 0 & \end{array}\right]_{8\times 8} \tag{237}$$

leaves all of the rows of $[\bar{E} \;\; \bar{G}]$ the same except zeroing the last two rows. Thus the sensitivities with respect to the parameter b_3 may be determined as linear combinations of x(t;u), $\xi^{(1)}(t;u)$, and $\xi^{(2)}(t;u)$, and six differential equations are now required to compute the complete sensitivity system. Finally, by Corollary 7.1 there exists no control which will completely zero X(t;u) in a finite interval of time.

4. PARAMETER IDENTIFIABILITY

The final topic considered in this section is identifiability of the unknown parameters in the system S_{LC}. One could specialize Lemma 4 and test the singularity of the p x p local identifiability matrix:

$$M = \int_0^{t_f} [v^{(i)}(t;u)/v^{(j)}(t;u)]dt \tag{238}$$

Identifiability would then be determined according to whether or not M has rank p. Such an approach is appropriate for the time-varying linear system, S_{TV}, and is the method commonly known in the literature (e.g., (Ref 102)). However, for the time-varying system, S_{LC}, the definition of identifiability itself (namely, V mapping R^p into the output space $Y = L_2(0,t;R^m)$ is one-to-one -- see Definition 5) may be used to obtain direct algebraic conditions for identifiability (both local and structural).

As discussed in Section II.7, the conditions for zero-input and zero-state identifiability are studied separately. This separate analysis allows one to determine whether a parameter is identifiable from either the transient response alone or the forced response alone or from both. Lemma 5 then concludes that the combined identifiable subspace is just the sum of the zero-input and zero-state identifiable subspaces.

Finally, we note the approach taken here in treating identifiability and the results obtained are new. Background literature on identifiability for the system S_{LC} was discussed in Section I.1.d. Although there have been previous structural and algebraic conditions for identifiability these previous approaches generally rely upon special canonical form representations of the system S_{LC}. As done throughout this research, we do not make any assumptions (except differentiability) concerning the

system matrices, and so the results may be applied either with or without such canonical forms.

a. Zero-Input Identifiability

To discuss zero-input identifiability it is convenient to rearrange equation (191) into the form

$$V_{z.i._j}(t) = [v^{(1)}_{z.i._j}(t) \ldots v^{(p)}_{z.i._j}(t)]_{1xp}$$

$$= a^T(t)E^*_j \tag{239}$$

$$V_{z.i.}(t) = \begin{bmatrix} V_{z.i._1}(t) \\ \cdot \\ \cdot \\ \cdot \\ V_{z.i._m}(t) \end{bmatrix} = \begin{bmatrix} a^T(t)E^*_1 \\ \cdot \\ \cdot \\ \cdot \\ a^T(t)E^*_m \end{bmatrix}_{mxp}$$

$$= a^T(t)*E^* \tag{240}$$

where we define

$$E^*_j \equiv [E^{(1)T}_j \ldots E^{(p)T}_j]_{2nxp} \tag{241}$$

and

$$E^* \equiv \begin{bmatrix} E^*_1 \\ \cdot \\ \cdot \\ \cdot \\ E^*_m \end{bmatrix}_{2nmxp} \tag{242}$$

Using this notation and the definitions of identifiability (Definitions 5, 6, and 7) we may directly obtain algebraic conditions for zero-input identifiability. We define the following zero-input identifiable subspaces:

$M_{z.i._j}$ ≡ zero-input identifiable subspace with respect to the output component $y_{z.i._j}$

$M_{z.i.}$ ≡ zero-input identifiable subspace

Using this notation we obtain the following central theorem:

Theorem 9

For the system S_{LC} it is true that:

i) $M_{z.i._j} = R(E_j^{*T})$ (243)

ii) $M_{z.i.} = R(E^{*T})$ (244)

iii) The system S_{LC} is zero-input identifiable if and only if the rank of E^* is p.

Proof

By Definition 5, the zero-input non-identifiable subspace relative to the output component $y_{z.i._j}(\cdot)$ is the null space of

$$V_{z.i._j}(\cdot) = a^T(\cdot)E_j^* \tag{245}$$

But since the scalar functions $a_j(\cdot)$ are linearly independent (and hence non-zero almost everywhere), the null space of $V_{z.i._j}(\cdot)$ is just the null space of E_j^*. The identifiable subspace is the orthogonal complement of

the non-identifiable subspace, which, by Lemma 1 is the range space of E_j^{*T}. Then S_{LC} is zero-input identifiable relative to $y_{z.i._j}$ if and only if the rank of E_j^* is p.

From equation (140), the conclusions of part ii) and iii) follow in a manner similar to that above.

Q.E.D.

Theorem 9 provides the central algebraic results on zero-input identifiability. The results are only dependent on the structure of the system matrices and initial condition vector; they are completely general; and no integrals are required to obtain the zero-input identifiability matrix. Notice the connection between this theorem on local identifiability and zero-input insensitivity; the matrix E^* used here has the same elements as the E matrix which was used in Theorem 5 in connection with zero-input insensitivity. If the system is totally zero-input insensitive (i.e., $E^* = 0$), then the non-identifiable subspace is all of R^p and there is no information with which to identify any of the parameters from the transient response. Such a connection between insensitivity and non-identifiability was noted by Bonivento (Ref 55), but no specific results on identifiability were obtained.

As would be expected, this theorem on zero-input identifiability reduces to previously known results concerning system observability (e.g., Ref 33)) for the special case in which d(b) = b is the only unknown system parameter vector:

Corollary 4.4.1

If d(b) = b is the only unknown system parameter vector, then S_{LC} is "observable" if and only if the rank of the n x mn matrix

$$[C^T \mid (CA)^T \mid \ldots \mid (CA^{n-1})^T]_{nxnm} \qquad (246)$$

is n.

Proof

Since d(b) = b is the only unknown parameter, it is true that

$$E_j^{(i)} = [(C_j e_i) \ldots (C_j A^{2n-1} e_i)] \tag{247}$$

where e_i is the i^{th} Euclidean basis vector. Then

$$E_j^* = \begin{bmatrix} C_j \\ C_j A \\ \vdots \\ C_j A^{2n-1} \end{bmatrix}_{2nxn} \tag{248}$$

Using the Cayley-Hamilton theorem (Ref 33) it is easily shown that the rank of

$$E^* = \begin{bmatrix} E_1^* \\ E_2^* \\ \vdots \\ E_m^* \end{bmatrix}_{2nmxn} \tag{249}$$

is equal to the rank of the matrix (246).

Q.E.D.

Finally, we note that if the system S_{LC} has eigenvalues with strictly negative real parts (S_{LC} is asymptotically stable), or if it has a single zero eigenvalue with zero sensitivities with respect to that eigenvalue, then all of the sensitivity operators will be asymptotically stable in the first case or marginally stable in the second. (See (Ref 52) or Corollary 2.4) Then although the scalar functions $a_j(\cdot)$, j = 1, 2, ...2n, are

linearly independent and hence non-zero almost everywhere, after a few system time constants they will be so near to zero that there is very little "signal strength" remaining in the sensitivities. Therefore, for all practical purposes, the zero-input response will no longer be identifiable in the steady-state.

b. Zero-State Identifiability

The question of zero-state identifiability is somewhat more complex than zero-input identifiability because the input function, $u(\cdot) \in L_2(0,t_f;R^r)$, can also affect identification capability. Certainly if $u(\cdot) = 0$, then $y_{z.s.}(\cdot;u) = 0$ and there is no information whatsoever from the zero-state response. The question of input design for the purpose of optimal parameter identification was discussed in Section II.8 and computational algorithms to calculate the optimal input are considered in Section VII. Here we derive necessary conditions for zero-state identifiability; sufficiency conditions are dependent upon the control input which one selects.

Like the previous section, it is helpful to rearrange equation (192) into the form

$$V_{z.s._j}(t;u) = [v^{(1)}_{z.s._j}(t;u) \cdots v^{(p)}_{z.s._j}(t;u)]_{1xp}$$

$$= g^T(t;u)G^*_j \tag{250}$$

$$V_{z.s.}(t;u) = \begin{bmatrix} V_{z.s._1}(t;u) \\ \vdots \\ V_{z.s._m}(t;u) \end{bmatrix} = \begin{bmatrix} g^T(t;u)G^*_1 \\ \vdots \\ g^T(t;u)G^*_m \end{bmatrix}$$

$$= g^T(t;u)*G^* \tag{251}$$

where we define

$$G_j^* = [G_j^{(1)^T} G_j^{(2)^T} \cdots G_j^{(p)^T}]_{2nrxp} \tag{252}$$

and

$$G^* = \begin{bmatrix} G_1^* \\ \cdot \\ \cdot \\ \cdot \\ G_m^* \end{bmatrix}_{2nrmxp} \tag{253}$$

Then defining the following zero-state identifiable subspaces:

$M_{z.s._j}(u) \equiv$ zero-state identifiable subspace relative to the output component $y_{z.s._j}(\cdot;u)$;

$M_{z.s.}(u) \equiv$ zero-state identifiable subspace;

we may use the above notation to obtain the central result on zero-state identifiability:

Theorem 10

For the system S_{LC} it is true that

i) $M_{z.s._j}(u) \subset R(G_j^{*T})$ (254)

ii) $M_{z.s.}(u) \subset R(G^{*T})$ (255)

iii) S_{LC} is zero-state identifiable only if the rank of G^* is p.

Proof

For part i), the non-identifiable subspace with respect to the output component $y_{z.s._j}(\cdot;u)$ is the null space of $V_{z.s._j}(\cdot;u)$. From equa-

tion (250) we see that

$$N(G_j^*) \subset N(V_{z.s._j}(\cdot;u)) \tag{256}$$

Hence

$$M_{z.s._j}(u) = N(V_{z.s._j}(\cdot;u))^{\perp}$$

$$\subset N(G_j^*)^{\perp} = R(G_j^{*T}) \tag{257}$$

where the last step follows from Lemma 1.

Part ii) is obtained in a similar manner. Then the system S_{LC} is zero-state identifiable if and only if $M_{z.s.}(u) = R^p$ which can be true only if rank G^* is p.

Q.E.D.

The conditions of Theorem 10 can be made sufficient as well as necessary (i.e., the subspace relations are replaced by equalities) provided that the 2nr scalar functions of $t \in [0,t_f]$

$$\int_0^t a_j(t-s)u_k(s)ds \qquad k = 1, 2, \ldots r;\ j = 1, 2, \ldots 2n; \tag{258}$$

are linearly independent. This condition may be tested by considering whether or not the 2nr x 2nr Gram matrix (see equations (185) - (186))

$$\int_0^{t_f} g(t;u)g^T(t;u)dt \tag{259}$$

is non-singular. Since the scalar functions $a_j(\cdot)$, $j = 1, 2, \ldots 2n$, are linearly independent on every interval of finite length, the condition of linear independence of the functions (258) is not too severe from a pure mathematical point of view. However, from a practical point of view, if the system has stable eigenvalues then the scalar functions $a_j(\cdot)$ all tend to zero in the steady state. This is easily seen by recalling

that the functions $a_j(t)$ are linearly independent combinations of the functions $t^j e^{\lambda_k t}$ (see Appendix A). Therefore, in the steady state the potential for "structural" zero-state identifiability can only be realized if the control input is selected to "excite" the modes of the system to a sufficient degree to maintain linear independence of the functions (258). This is the problem referred to by Åström and Eykhoff (Ref 80) as the requirement for "persistent excitation". For the class of sinusoidal inputs this problem has been studied by Mehra (Ref 98), Kim and Lindorff (Ref 95), Hoberock and Stewart (Ref 94), and others to determine sufficient numbers of input frequencies to ensure steady-state identifiability.

Finally, we comment that the results of Theorem 10 may be applied to one column vector of B at a time. and thereby we may obtain the zero-state identifiable subspace relative to each control input component, $u_k(\cdot)$, k = 1, 2, ...r. Then by selecting all of the controls to be identically zero except u_k, the parameter components which lie in the non-identifiable subspace of this input will not, to first order, influence the observed output. Again this may be a possible systematic way to reduce the number of influencing parameters in a series of parameter identification experiments. Indeed, the non-zero input component, u_k, may even be selected to optimize the estimation capability for those parameters which lie in its identifiable subspace. (See Section VII)

c. Summary of Identifiability Conditions for S_{LC}

In the previous two subsections the zero-input and zero-state local identifiable subspaces, $M_{z.i.}$ and $M_{z.s.}$, have been determined. By using the operator definitions of local identifiability (Definition 5) and the algebraic representation of the sensitivity operators for the system S_{LC},

(Theorem 4), both of these subspaces may be obtained through purely algebraic means which involve neither the determination of the sensitivity operators themselves nor integration for determining the local identifiability matrices

$$M_{z.i.} = \int_0^{t_f} [v^{(i)}_{z.i.}(t)/v^{(j)}_{z.i.}(t)]dt \tag{260}$$

$$M_{z.s.} = \int_0^{t_f} [v^{(i)}_{z.s.}(t;u)/v^{(j)}_{z.s.}(t;u)]dt \tag{261}$$

These results and the approach taken are new and should offer useful information for practioners of system identification.

Finally, Lemma 8 says that the total identifiable subspace is

$$M = M_{z.i.} + M_{z.s.} \tag{262}$$

while the total non-identifiable subspace is

$$M^{\perp} = M^{\perp}_{z.i.} \cap M^{\perp}_{z.s.} \tag{263}$$

The system S_{LC} is then identifiable if and only if $M = R^p$.

Example

Again consider the second order system introduced in the example of Section IV.3, equations (229) (230). For simplicity we will examine only zero-state identifiability (we will assume the system (229) (230) has zero initial conditions). Using equation (231) we obtain:

$$G_1^* = [G_1^{(1)^T} \quad G_1^{(2)^T} \quad G_1^{(3)^T}]$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & b_3 & b_2 \\ b_3 & 2b_2b_3 & b_1 + b_2^2 \end{bmatrix} \tag{264}$$

and

$$G_2^* = [G_2^{(1)^T} \quad G_2^{(2)^T} \quad G_2^{(3)^T}]$$

$$= \begin{bmatrix} 0 & 0 & 1 \\ 0 & b_3 & b_2 \\ b_3 & 2b_2b_3 & b_1 + b_2^2 \\ 2b_2b_3 & 2b_1b_3 + 3b_2^2b_3 & 2b_1b_2 + b_2^3 \end{bmatrix} \tag{265}$$

Hence if b_3 is non-zero, then the system (229) - (230) is zero-state identifiable from either the output y_1 or the output y_2 (since the rank of G_1^* and G_2^* are both 3). However, if b_3 is zero, then only b_3 is zero-state identifiable from either output.

5. SUMMARY

In this section the geometric properties of Theorem 4 have provided algebraic conditions for insensitivity, sensitivity controllability, and identifiability for the time-invariant system S_{LC}. Some of these conditions are restatements or extensions of previously known results (e.g., conditions ror zero-input and zero-state insensitivity (Theorems 5 and 6), general algebraic conditions for sensitivity controllability, (Theorem 7), and definitive algebraic conditions for zero terminal sensitivity

(Corollary 7.1)). However, in other cases the conclusions and the approach taken are new (general algebraic conditions for local and structural identifiability (Theorems 9 and 10), and a systematic way to reduce the number of influencing components in a series of identification experiments (Corollary 5.1)). In all cases the operator approach of Section II provides a straight-forward means to derive the results, and no transformation of the system to a suitable canonical form is required for the analysis. A simple second-order system was used to illustrate the various algebraic conditions and to demonstrate the way in which these algebraic conditions can be used to provide structural as well as local information.

In the next three sections we concentrate on the computational aspects of Theorem 4 by considering parameter identification algorithms, control design for minimum sensitivity, and sensitivity operator design for improved parameter identification. Once again the operator approach of Section II is used as the basis for this discussion.

Section V

QUASILINEARIZATION PARAMETER IDENTIFICATION FOR LINEAR TIME-INVARIANT SYSTEMS, S_{LC}

The previous section investigated geometrical properties of the matrix-operator representation for parameter sensitivities in the linear time-invariant system, S_{LC}, (see Theorem 4). In this section the computational aspects of this representation are considered by using it as the basis of a quasilinearization algorithm for the identification of unknown parameters in the system S_{LC}. Quasilinearization is a well-known method for the estimation of parameters in such systems (e.g., Kumar and Sridhar (Ref 86)), and so the unique element of this section is to illustrate how the results of Theorem 4 and Appendix A may be used as an alternative to differential equation computation of the parameter sensitivities.

The quasilinearization algorithm utilized is based upon the discussion of Section II.6. We noted in that section that the computational requirements of quasilinearization are fairly representative of gradient and Newton-Raphson algorithms as well. However, the Newton-Raphson method requires one to compute the second order as well as the first order parameter sensitivities. Appendix B presents a matrix-operator form of the second order sensitivities which is equivalent to the one developed in Section III and used here for the first order parameter sensitivities.

1. QUASILINEARIZATION ALGORITHM

The measured system output is denoted, $y(t)$, $t \in [0,t_f]$, and the model system output at the initial guess of the parameter vector,

$\hat{b}_o \in R^p$, is designated $y(t;\hat{b}_o)$, $t \in [0,t_f]$. The output error cost functional which we wish to minimize is (see Section II.6.d):

$$J_Q(\hat{b}) = \int_0^{t_f} [(y(t) - y(t;\hat{b}))/Q(t)(y(t) - y(t;\hat{b}))]dt \tag{266}$$

where the $Q(t) \in R^{mxm}$ is assumed to be symmetric and uniformly positive definite for all $t \in [0,t_f]$. From Section II.6.b and II.6.d the quasi-linearization algorithm to minimize this cost functional is obtained from the sequence of steps:

$$\Delta b_j = (V^*QV)_j^{-1} V^*Qz_j \tag{267}$$

$$\hat{b}_{j+1} = \hat{b}_j + \Delta b_j \tag{268}$$

where $(V^*QV)_j^{-1}$ is the p x p local "information" matrix[1]

$$(V^*QV)_j = \int_0^{t_f} [v^{(i)}(t;\hat{b}_j)/Q(t)v^{(k)}(t;\hat{b}_j)]dt \quad ; \tag{269}$$

V^*Qz_j is one-half the negative gradient of $J_Q(\hat{b}_j)$

$$V^*Qz_j = \int_0^{t_f} [v^{(i)}(t;\hat{b}_j)/Q(t)z(t;\hat{b}_j)]dt \quad ; \tag{270}$$

and $z(t;\hat{b}_j) \equiv y(t) - y(t;\hat{b}_j)$ is the output error at the current estimate $\hat{b}_j \in R^p$. Recall from Section II.7 that the local information matrix, $(V^*QV)_j$, possesses a bounded inverse if the system is locally identifiable at $\hat{b}_1 \in R^p$. We will assume that the system is locally identifiable along

---

[1]Note that if the output is only sampled at discrete points, then the integrals in equations (269) and (270) would be replaced by summations over these discrete time points.

the sequence of estimates, $\hat{b}_j$, and hence that these inverses do indeed exist.

Notice that if a direct "sensitivity system" differential equation approach were used in determining the output, $y(t;\hat{b}_j)$, and sensitivities $v^{(i)}(t;\hat{b}_j)$, $i = 1, 2, \ldots p$, then each iteration would require the solution of possibly $n(p + 1)$ coupled linear differential equations[1] in addition to the $(p^2 + 3p)/2$ quadrature integrals required for computing $(V^*QV)_j$ and V^*Qz_j. This number of differential equations and integrals may, of course, be reduced if the controllability properties of the sensitivity system are taken into account (see Section IV.3), and Gupta and Mehra (Ref 84) present the details of this as applied to parameter identification algorithms. However, as a computational alternative, we apply the results of Section III and Appendix A, thereby replacing the solution of the differential equations with the determination of the eigenvalues of $A(\hat{b}_j)$ and the solution of, at most, $2nr$ quadrature integrals for computing the parameter sensitivities[1].

The computational details of this method are presented in the subsequent algorithm. The notation used is based upon the development of Appendix A in which the matrix-operator representation of the parameter sensitivities is factored directly into "component" form utilizing the basis functions, $t^j e^{q_k t}$, for real eigenvalue q_k (or $t^j e^{\sigma_k t} \cos\omega_k t$ and $t^j e^{\sigma_k t} \sin\omega_k t$ for complex eigenvalues) rather than the linear combinations of the scalar functions, $a_j(t)$ (see Theorem 4). In particular, the system nominal output and output sensitivities are computed from the relations

$$y(t;\hat{b}_j) = F^{(0)} f(t) + H^{(0)} \int_0^t f(t - s)^* u(s) ds \qquad (271)$$

---

[1]Recall that r is the control dimension, n is the state dimension, and p is the parameter dimension.

$$v^{(i)}(t;\hat{b}_j) = F^{(i)}f(t) + H^{(i)} \int_0^t f(t - s)*u(s)ds \qquad (272)$$

$i = 1, 2, \ldots p$

where $f(\cdot)$ is a 2n-dimensioned vector function defined in Appendix A and which has elements of the form $t^j e^{q_k t}$; $F^{(0)}$, $F^{(i)}$, $H^{(0)}$, and $H^{(i)}$, $i = 1, 2, \ldots p$, are m x 2n and m x 2nr dimensioned matrices and are computed as described in Appendix A; and we define the special product

$$f(t - s)*u(s) \equiv \begin{bmatrix} f_1(t - s)u(s) \\ f_2(t - s)u(s) \\ \vdots \\ f_{2n}(t - s)u(s) \end{bmatrix}_{2nrx1} \qquad (273)$$

This form is used for computation rather than the one used in Theorem 4 and Section IV because the convolutions of equations (271) - (272) may be directly transformed into the time-varying linear combination of quadrature integrals through relations such as

$$(t - s)e^{q_k(t - s)} = te^{q_k t}e^{-q_k s} - e^{q_k t}se^{-q_k s} \qquad (274)$$

However, either form is equivalent and the differences are merely a matter of bookkeeping on the digital computer. The form (271) - (272) will also be utilized in the computational applications of Sections VI and VII. The algorithm for quasilinearization parameter identification is now stated:

Quasilinearization Algorithm to Identify

the Parameters in the System S_{LC}

1) Based upon the current estimate of the unknown parameters, $\hat{b}_j \in R^p$, determine the distinct eigenvalues of $A(\hat{b}_j)$, q_k, $k = 1, 2, \ldots \rho$, and their multiplicity, n_k, in the characteristic polynomial of $A(\hat{b}_j)$. Note that if the eigenvalues of $A(\hat{b}_{j-1})$ have been computed, then iterative techniques for computing the eigenvalues of $A(\hat{b}_j)$ might be useful. (e.g., (Ref 11: 455 - 495)). Indeed, by using the expression (157) suggested in Section III for computing the eigenvalues sensitivities, a first guess of the new eigenvalues might be computed from the Taylor's series relation

$$q_k^j = q_k^{j-1} + \sum_{i=1}^{p} (q_k)_{(i)} (\hat{b}_i^j - \hat{b}_i^{j-1}) \tag{275}$$

where the superscripts "j" and "j-1" refer to the j^{th} and $j-1^{th}$ iterations, respectively. (See the example of Section V.2 for illustration.)

2) Using the eigenvalues of $A(\hat{b}_j)$, invert the 2n x 2n Vandermonde matrix, Λ. (See Appendix A.) Then using the elements of Λ^{-1}, and the input quantities A, B, C, d, $A_{(i)}$, $B_{(i)}$, $C_{(i)}$, and $d_{(i)}$, $i = 1, 2, \ldots p$, compute the matrices

$$\bar{\bar{F}} = \begin{bmatrix} F^{(0)} \\ F^{(1)} \\ \vdots \\ F^{(p)} \end{bmatrix}_{m(p+1)x2n} \qquad \bar{\bar{H}} = \begin{bmatrix} H^{(0)} \\ H^{(1)} \\ \vdots \\ H^{(p)} \end{bmatrix}_{m(p+1)x2nr} \tag{276}$$

(See Appendix A for the definition and computation of the above quantities.)

3) Compute and place in temporary storage the (p + 1) m-dimensional vector functions of $t \in [0,t_f]$, $y(t;\hat{b}_j)$, $v^{(i)}(t;\hat{b}_j)$, $i = 1, 2, \ldots p$. (See equations (271) - (272).)

4) Compute

$$\Delta b_j = (V^*QV)_j^{-1} V^*Qz_j \tag{277}$$

using equations (269) - (270). Note that the "controllability" properties of the sensitivities may be used to reduce the number of integrals (summations) required in computing the matrix $(V^*QV)_j$ and the vector V^*Qz_j (see Section IV.3). In particular, if ν is the dimension of the mp x 2n(1 + r) dimensioned matrix

$$[F \quad G] = \begin{bmatrix} F^{(1)} & G^{(1)} \\ F^{(2)} & G^{(2)} \\ \vdots & \vdots \\ F^{(p)} & G^{(p)} \end{bmatrix} \tag{278}$$

then there are only ν linearly independent functions from the mp functions $\{v^{(i)}(\cdot;\hat{b}_j);\ i = 1, 2, \ldots p\}$. Therefore, there will only be $(\nu + 1)\nu/2$ integrals (summations) required to compute the symmetric matrix $(V^*QV)_j$ and ν to compute the vector V^*Qz_j.

5) Set

$$\hat{b}_{j+1} = \hat{b}_j + \Delta b_j \tag{279}$$

and either return to 1) or stop when $||y - y(\hat{b}_j)||$ is less than some specified small value.

Stop.

2. EXAMPLE

Once again consider the second order system introduced in the examples of Sections IV.2 - IV.4:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ b_3 \end{bmatrix} u(t) \qquad t \in [0,t_f] \qquad (280)$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \qquad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad (281)$$

The nominal value of $b \in R^3$ is again assumed to be $\hat{b}_o = [-2 \quad -3 \quad 1]^T$, and so the nominal A matrix has eigenvalues $q_1 = -1$ and $q_2 = -2$.

To use the algorithm of the previous subsection, we must first determine the matrices $F^{(0)}$, $F^{(i)}$, $H^{(0)}$, $H^{(i)}$, $i = 1, 2, 3$, as described in Appendix A. To do so requires that we invert the Vandermonde matrix

$$\Lambda = \begin{bmatrix} 1 & q_1 & q_1^2 & q_1^3 \\ 0 & 1 & 2q_1 & 3q_1^2 \\ 1 & q_2 & q_2 & q_2^3 \\ 0 & 1 & 2q_2 & 3q_2^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 1 & -2 & 4 & -8 \\ 0 & 1 & -4 & 12 \end{bmatrix} \qquad (282)$$

This inverse is easily shown to be

$$\Lambda^{-1} = \begin{bmatrix} -4 & 4 & 5 & 2 \\ -12 & 8 & 12 & 5 \\ -9 & 5 & 9 & 4 \\ -2 & 1 & 2 & 1 \end{bmatrix} \tag{283}$$

Then since there is a single control input we may determine $\bar{F}$ and $\bar{H}$ from the equations

$$\bar{F} = \begin{bmatrix} F^{(0)} \\ F^{(1)} \\ F^{(2)} \\ F^{(3)} \end{bmatrix} = \bar{E}\Lambda^{-1} \qquad \bar{H} = \begin{bmatrix} H^{(0)} \\ H^{(1)} \\ H^{(2)} \\ H^{(3)} \end{bmatrix} = \bar{G}\Lambda^{-1} \tag{284}$$

where $\bar{E}$ and $\bar{G}$ are given by expression (236) of Section IV.3. Thus, for these nominal parameter values, the matrices $\bar{F}$ and $\bar{H}$ become

$$\bar{F} = \begin{bmatrix} 2 & 0 & -1 & 0 \\ -2 & 0 & 2 & 0 \\ -3 & 2 & 3 & 1 \\ 5 & -2 & -5 & -2 \\ 4 & -2 & -4 & -2 \\ -6 & 2 & 6 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \bar{H} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 \\ -2 & 1 & 2 & 1 \\ 3 & -1 & -3 & -2 \\ 3 & -1 & -3 & -2 \\ -4 & 1 & 4 & 4 \\ 1 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 \end{bmatrix} \tag{285}$$

Then the output and output sensitivities are computed from the expressions (271) and (272) where

$$f(t) = [e^{-t} \quad te^{-t} \quad e^{-2t} \quad te^{-2t}]^T \tag{286}$$

Notice that the integral $\int_0^t f(t-s)u(s)ds$ may be obtained as time-varying

combinations of the four quadrature integrals:

$$\int_0^t e^s u(s)ds; \quad \int_0^t se^s u(s)ds; \quad \int_0^t e^{2s} u(s)ds; \quad \int_0^t se^{2s} u(s)ds \tag{287}$$

Hence, there are only four integrals required to compute the complete set of parameter sensitivities for each iteration of the quasilinearization algorithm. It is interesting to note that if a sensitivity system differential equation approach were utilized in computing the sensitivities, there would be either eight differential equations for a direct solution, or a minimum of six if reduction techniques were applied (see the example of Section IV.3). Thus, in either case the number of required differential equations would be greater than the number of required quadratures by the operator method.

Once the new estimate $\hat{b}_1$ has been computed from equation (279), to perform another iteration of the algorithm requires a recomputation of the eigenvalues of $A(\hat{b}_1)$. A first guess of the new eigenvalues can be obtained by using the eigenvalue sensitivities computed from a Taylor's series expansion about $\hat{b}_o$. For this example, the new guess would be

$$q_1 \simeq q_1^o + (q_1)_{(1)} \Delta b_1^o + (q_1)_{(2)} \Delta b_2^o \tag{288}$$

$$q_2 \simeq q_2^o + (q_2)_{(1)} \Delta b_1^o + (q_2)_{(2)} \Delta b_2^o \tag{289}$$

where the superscript "o" here denotes the eigenvalue and parameter components at the initial guess $\hat{b}_o$. The eigenvalue sensitivities may be computed on the basis of equation (157) of Section III. In fact, since the C matrix is independent of any of the parameter components which appear in the A matrix, we may use the column vectors of $\overline{F}$ to determine these eigenvalue sensitivities (the column vectors of $\overline{H}$ could alternately

be utilized since the parameters of A do not appear in the B matrix either). Partitioning $\overline{F}$ into the form

$$\overline{F} = \begin{bmatrix} F^{(0)}_{(1)} & F^{(0)}_{(2)} & F^{(0)}_{(3)} & F^{(0)}_{(4)} \\ \vdots & & & \vdots \\ F^{(3)}_{(1)} & \cdots & & F^{(3)}_{(4)} \end{bmatrix}_{8x4} \tag{290}$$

we may compute these eigenvalue sensitivities from the relations:

$$F^{(i)}_{(2)} = (q_1)_{(i)} F^{(0)}_{(1)} \tag{291}$$

$$F^{(i)}_{(4)} = (q_2)_{(i)} F^{(0)}_{(3)} \tag{292}$$

Hence

$$(q_1)_{(1)} = -1 \qquad (q_1)_{(2)} = -1 \qquad (q_1)_{(3)} = 0 \tag{293}$$

$$(q_2)_{(1)} = -1 \qquad (q_2)_{(2)} = 2 \qquad (q_2)_{(3)} = 0 \tag{294}$$

Thus the eigenvalue sensitivities may be obtained in a natural way from the other computations which take place in computing the parameter sensitivities via the operator method.

3. SUMMARY

In this section we have used the representation of Section III to obtain a new computational algorithm for quasilinearization parameter identification in linear time-invariant ordinary differential equation systems. This is one of the more direct computational applications and it shows how the sensitivity system differential equations may be replaced by the solution of a relatively low number of quadrature integrals.

Furthermore, the methods for computing the eigenvalue sensitivities from this form (expression (157), Section III) and the use of the "controllability" properties (Section IV.3) have been discussed in relation to reducing the total computational requirements for this approach.

In Sections VI and VII we consider further computational applications by applying the results of Section III to minimum sensitivity control design and sensitivity operator design for optimal parameter identification.

Section VI

MINIMUM SENSITIVITY CONTROL DESIGN FOR THE LINEAR TIME-INVARIANT SYSTEM S_{LC}

In Section II.6 a gradient method for open-loop minimum sensitivity control design was developed for the Hilbert space linear system S_H. In this section we apply this method to the linear time-invariant ordinary differential equation system, S_{LC}, and consider this as a computational alternative to previously employed Riccati equation techniques (e.g., Kahne (Ref 70)). Like the previous section, this illustrates the computational features of the matrix-operator form of the parameter sensitivities which is presented in Section III and Appendix A.

1. GRADIENT METHOD OF SOLUTION

To illustrate some important computational aspects, the cost functional of Section II.5 is generalized slightly to consider minimization of

$$J_S(u) = [Y(t_f;u)/S_f Y(t_f;u)] + \int_0^{t_f} [Y(t;u)/S(t)Y(t;u)]dt + \int_0^{t_f} [u(t)/O(t)u(t)]dt \qquad (295)$$

where $Y(t;u) \in R^{m(p+1)}$ is the augmented vector

$$Y(t;u) \equiv \begin{bmatrix} y(t;u) \\ v^{(1)}(t;u) \\ \vdots \\ v^{(p)}(t;u) \end{bmatrix} \qquad (296)$$

The real matrices S_f and $S(t)$ are assumed to be nonnegative definite and symmetric, and the real matrix $O(t)$ is assumed to be symmetric and uniformly positive definite. The latter assumption is required to ensure existence of a unique minimizing control. We note that the cost functional $J_S(u)$ is quite general as it weights the sensitivities along the nominal trajectory as well as at the terminal time t_f. Such a cost functional has been used by a number of other researchers in the minimum sensitivity control design problem (e.g., Kahne (Ref 70)), and the weighting matrices may be manipulated to achieve various design objectives.

Following the approach of Section II.5, we may separate $Y(t;u)$ into its zero-input portion, $Y_{z.i.}(t)$, and zero-state portion, $Y_{z.s.}(t;u)$. Then using the notation of the previous section which is defined in detail in Appendix A, these may be computed from the matrix-operator relations

$$Y_{z.i.}(t) = \overline{F}\, f(t) \tag{297}$$

$$Y_{z.s.}(t;u) = \overline{G} \int_0^{t_f} f(t - s)u(s)ds. \tag{298}$$

For notational simplicity we will assume in this Section that $u(\cdot) \in L_2(0,t_f;R)$ is a single control input (that is, $r = 1$). The results are easily extended to the multi-input case. By substituting equations (297) - (298) into (295) and performing an interchange of the order of integration, the cost functional $J_S(u)$, equation (295), may be put into the form

$$J_S(u) = k + \int_0^{t_f} [u(t)/2h(t) + P(t)u]dt \tag{299}$$

where

$$k = [Y_{z.i.}(t_f)/S_f Y_{z.i.}(t_f)] + \int_0^{t_f} [Y_{z.i.}(t)/S(t)Y_{z.i.}(t)]dt \qquad (300)$$

$$h(t) = f^T(t_f - t)\bar{H}^T S_f Y_{z.i.}(t_f)$$

$$+ \int_t^{t_f} f^T(s - t)\bar{H}^T S(s)Y_{z.i.}(s)ds \qquad (301)$$

$$P(t)u = 0(t)u(t) + f^T(t_f - t)\bar{H}^T S_f Y_{z.s.}(t_f;u)$$

$$+ \int_t^{t_f} f^T(s - t)\bar{H}^T S(s)Y_{z.s.}(s;u)ds \qquad (302)$$

The quadratic cost functional $J_S(u)$ may then be minimized by gradient (or conjugate gradient) iterations where the gradient of $J_S(u)$ is

$$\nabla J_S(t;u) = 2(h(t) + P(t)u) \qquad (303)$$

$$t \in [0,t_f]$$

(See, e.g., Luenberger (Ref 20).)

For such gradient minimization techniques, the gradient function must be computed for each new guess of the optimal control. Therefore, a major element of the computation time for such an approach will be the computation of the gradient function itself, equation (303). Therefore, we discuss the computational features associated with computing this gradient function from equations (301) - (303).

The function $h(\cdot)$ must be computed once and stored. The terminal sensitivity term (the first term of equation (301)) involves no integrals but merely matrix manipulations and the determination of the 2n-dimensioned vector function $f(\cdot)$ from the eigenvalues of A (see Appendix A for details). The trajectory sensitivity term (the second term of (301)) requires 2n

integrals for each separate control input. It is most important to point out that these indicated convolutions may be transformed into quadratures by the methods discussed in Section V (see equation (274)).

Next, the function $P(\cdot)u$ must be recomputed for each new guess of the control input. An inherent part of the computation is obtaining $Y_{z.s.}(\cdot;u)$ from equation (298). As discussed in Section V, this also requires the solution of 2n quadrature integrals for each separate control input. Once computing $Y_{z.s.}(\cdot;u)$ the terminal sensitivity term of $P(\cdot)u$ (the second term of (302)) requires only matrix operations and no additional integrals. However, the trajectory sensitivity term (the third term of (302)) again requires the solution of 2n integrals for each control input. Again it is significant to note that the indicated convolutions may be transformed into quadratures.

To summarize, the computational requirements for computing the gradient function, equation (303), are quite minimal and increase only linearly with state dimension n and control dimension r; the computational requirements are virtually independent of the parameter dimension p. It is also significant to point out that the terminal sensitivity terms inherently do not require any convolutions to be solved; therefore, if the trajectory sensitivity weighting, $S(\cdot)$, is identically zero, then the gradient function requires no convolutions to be obtained and the gradient method may readily be applied to other more general linear systems (say, e.g., the linear time-varying system of Section III.1 where Theorem 1 is used to compute the partial derivatives of the state transition matrix). However, if $S(\cdot)$ is non-zero then the gradient function $\nabla J_S(\cdot;u)$ requires convolutions to be computed (see the second and third terms respectively of equations (301) and (302)), and, except for the

case of the linear time-invariant plant matrix, these convolutions cannot generally be transformed into the far more desirable quadrature integral form. Therefore, for the general case with this operator-gradient approach the computational differences between the terminal sensitivity problem and the trajectory sensitivity problem are quite large. This is not true of the Riccati equation exact method of solution examined in the next section, for there the computational requirements are virtually the same for both the trajectory sensitivity and terminal sensitivity cases.

Another important factor in the computational requirements of the gradient method of solution is the accuracy of an initial guess of an optimal control. The initial guess

$$u_o(t) = -0^{-1}(t)h(t) \tag{304}$$

might be selected for this control minimizes the functional

$$[u(t)/2h(t) + 0(t)u(t)] \tag{305}$$

and this is an approximation to the sensitivity cost functional $J_S(u)$, equation (299). However, if the optimal control for one set of conditions has been obtained and these conditions are perturbed slightly (say, for example, if a new set of initial conditions are selected or if linearization about a new nominal trajectory in a quasilinearization nonlinear system optimization is desired), then the old optimal control may be used as the first guess with the new conditions. Since the quadratic functional $J_S(u)$ is continuous, if the new conditions are not perturbed greatly from the old, the new optimal control will be "close" to the old optimal and convergence should be quite rapid. In general, this is an advantage of iterative solution techniques over exact methods

of solution in that prior information such as this may be utilized to reduce computational requirements.

Finally, we comment that the gradient type approach may be conveniently applied to a number of variations to the basic quadratic minimum sensitivity control problem which has been considered here. For instance, if it is required that the system output at the terminal, $y(t_f;u)$, meet a terminal equality constraint[1]

$$y(t_f;u) = F^{(0)}f(t_f) + G^{(0)} \int_0^{t_f} f(t_f - t)u(t)dt$$

$$= y_f \tag{306}$$

and if $G^{(0)}$ has rank m (i.e., the system is output controllable), then the gradient projection method of Rosen (Ref 24) may be easily applied to solve the terminal equality constraint minimum sensitivity control problem. (See Luenberger (Ref 20: 297 - 299) for the computational details of this method.) Since the projection operator

$$\pi = I - f^T(t_f - t)G^{(0)^T}(G^{(0)}M_c G^{(0)^T})^{-1} G^{(0)} \int_0^{t_f} f(t_f - t)(\cdot)dt \tag{307}$$

$$M_c = \int_0^{t_f} f(t_f - t)f^T(t_f - t)dt \tag{308}$$

used in this method is easily calculated, there is very little computational increase for solving this terminal equality constraint problem.

3. RICCATI EQUATION SOLUTION

In the previous section a gradient approach to minimizing the quad-

---

[1]See Appendix A for the definition of the m x 2n and m x 2nr dimensioned matrices $F^{(0)}$ and $G^{(0)}$.

ratic sensitivity cost functional, $J_S(u)$, equation (295), was examined. The results of Section III and Appendix A were applied and some highly efficient computational features were obtained for the time-invariant system S_{LC}. However, often an explicit solution to the open-loop minimum sensitivity control problem would be desired rather than the iterative methods of the previous section. In such a case a Riccati equation method of solution is a well-known technique for calculating the open-loop minimum sensitivity control law in linear systems. (See, e.g., Kahne (Ref 70).) We emphasize the fact that the control law determined in this manner is only the open-loop solution and cannot be made closed-loop (except by approximation - see, e.g., Lamont and Kahne (Ref 72)) as is the case for a Riccati equation solution when sensitivity is not included in the cost functional. This fact is well-known (e.g., Price and Deyst (Ref 75)) and stems from the fact that the optimal control is computed from a linear combination of both the optimal nominal state and state sensitivities (see equation (312)). However, the sensitivities are computed under the assumption that the control is open-loop, and so the form becomes mathematically incorrect when used in a closed-loop (feedback) manner.

Using the time-invariant sensitivity system of Section III

$$\dot{X}(t;u) = \bar{A}X(t;u) + \bar{B}\,u(t) \qquad X(0) = \bar{d} \qquad (309)$$

$$Y(t;u) = \bar{C}X(t;u) \qquad t \in [0,t_f] \qquad (310)$$

(see equations (92) - (94)), the cost functional $J_S(u)$, equation (295), may be written as

$$J_s(u) = [X(t_f;u)/\bar{C}^T S_f \bar{C} X(t_f;u)]$$

$$+ \int_0^{t_f} [X(t;u)/\bar{C}^T S(t) \bar{C} X(t;u)]dt$$

$$+ \int_0^{t_f} [u(t)/O(t)u(t)]dt \qquad (311)$$

In this form the minimum sensitivity control problem is a standard linear system quadratic cost functional optimal control problem which has the well-known solution given by (e.g., Bryson and Ho (Ref 5)):

$$u^*(t) = -(O(t))^{-1}\bar{B}^T\bar{R}(t)X^*(t) \qquad (312)$$

where $\bar{R}(t)$, $t \in [0,t_f]$ is obtained via the unique solution of the $n(p+1) \times n(p+1)$ matrix Riccati differential equation

$$\dot{\bar{R}}(t) = -\bar{R}(t)\bar{A} - \bar{A}^T \bar{R}(t) + \bar{R}(t)\bar{B}\, O^{-1}(t)\bar{B}^T \bar{R}(t) - \bar{C}^T S(t)\bar{C} \qquad (313)$$

with terminal boundary condition

$$\bar{R}(t_f) = \bar{C}^T S_f \bar{C}. \qquad (314)$$

The $n(p+1)$ augmented vector function $X^*(\cdot)$ is computed by integrating forward the sensitivity system differential equations (309) using the optimal control computed from (312).

Now let us review the computational requirements of this Riccati equation method of solution. Without reducing the dimension of $X(\cdot;u)$ (see Section IV.3) the symmetric Riccati matrix, $\bar{R}(\cdot)$, will have dimension $n(p+1) \times n(p+1)$ and so the number of differential equations increases in rough proportion to the square of n and the square of p. However, if the methods of Section IV.3 (or other approaches to minimal

order sensitivity models -- see (Refs 34 - 42)) are applied, then the dimension of $X(\cdot;u)$ may be reduced to $\gamma_1 \leq 2n(r + 1)$ (see equations (226) - (228)). Therefore, for linear time-invariant systems the computations for the Riccati equation approach can be made to increase only with the squares of n and r and be independent of the parameter dimension p. This is in contrast to the gradient method of solution where the amount of computation increases only linearly with n and r and is also independent of the parameter dimension p. Thus for systems with large state dimension, n, the gradient approach offers the potential of considerable computational reduction.

Finally, we note that the Riccati method of solution may also be utilized in cases in which a terminal output condition $y(t_f;u) = y_f$ must be achieved (see, e.g., Bryson and Ho (Ref 5: 158 - 163)). Bryson and Ho describe the method in considerable detail; therefore, we do not repeat these details here, but merely comment that an additional linear matrix differential equation of dimension $n(p + 1) \times n$ (equation (5.3.35)(Ref 5: 161)) must be solved in using such a technique.

In the next subsection we consider a simple second order example to help illustrate the computational features of the operator-gradient method and also point out the computational differences between iterative gradient method and the Riccati equation exact method of solution.

3. EXAMPLE

Once again we consider the second order system

$$\begin{bmatrix} \dot{x}_1(t;u) \\ \dot{x}_2(t;u) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \end{bmatrix} + \begin{bmatrix} 0 \\ b_3 \end{bmatrix} u(t) \qquad (315)$$

$$\begin{bmatrix} x_1(0;u) \\ x_2(0;u) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad t \in [0, 1] \tag{316}$$

$$\begin{bmatrix} y_1(t;u) \\ y_2(t;u) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \end{bmatrix} \tag{317}$$

with nominal parameter vector $b_o = [-2 \quad -3 \quad 1]^T$. Since this example was considered in Sections IV.2, IV.3, IV.4, and V.2, we will use many of the results from those sections to shorten the discussion here.

As described in Section V.2 we may write the zero-input and zero-state augmented response in the form

$$Y_{z.i.}(t) = \bar{F}\, f(t) \qquad Y_{z.s.}(t;u) = \bar{H} \int_0^t f(t - s)u(s)ds \tag{318}$$

where $\bar{F}$, $\bar{H}$, and $f(t)$ are given by equations (285) - (286). Note that only four quadrature integrals, equation (287), are required in computing $Y_{z.s.}(t;u)$, $t \in [0, 1]$.

We will consider minimizing both a terminal sensitivity cost functional

$$J_1(u) = \int_0^1 u^2(t)dt + c_1[Y(1;u)/Y(1;u)] \tag{319}$$

and the trajectory sensitivity cost functional

$$J_2(u) = J_1(u) + c_2 \int_0^1 [Y(t;u)/Y(t;u)]dt \tag{320}$$

where the scalar constants c_1 and c_2 may be assigned various weighting values. The gradient of the terminal sensitivity cost functional may then be computed by

$$\nabla J_1(t;u) = 2(h_1(t) + P_1(t)u) \tag{321}$$

where

$$h_1(t) = c_1 f^T(1 - t)\bar{H}^T Y_{z.i.}(1) \tag{322}$$

$$P_1(t)u = u(t) + c_1 f^*(1 - t)\bar{H}^T Y_{z.s.}(1;u) \tag{323}$$

$$f^T(1 - t) = \left[e^{t-1} \quad (1-t)e^{t-1} \quad e^{2(t-1)} \quad (1-t)e^{2(t-1)} \right] \tag{324}$$

Notice that the only integrals required to compute the $P_1(t)u$ term are the four quadratures required in computing $Y_{z.s.}(t_f;u)$ from equation (318).

Next consider the gradient of the trajectory sensitivity cost functional, equation (320). This may be computed from

$$\nabla J_2(t;u) = 2(h_2(t) + P_2(t)u) \tag{325}$$

where

$$h_2(t) = h_1(t) + c_2 \int_t^1 f^T(s - t)\bar{H}^T Y_{z.i.}(s)ds \tag{326}$$

$$P_2(t)u = P_1(t)u + c_2 \int_t^1 f^T(s - t)\bar{H}^T Y_{z.s.}(s;u)ds \tag{327}$$

Note that we may work with $f^T(s - t)$ in the factored form

$$f^T(s - t) = \left[e^t e^{-s} \quad e^t(se^{-s} - te^{-s}) \quad e^{2t}e^{-2s} \quad e^{2t}(se^{-2s} - te^{-2s}) \right] \tag{328}$$

in order to transform the four convolutions of equations (326) and (327) into four quadratures each. Thus, for this trajectory sensitivity gradient there are a total of four quadratures to compute the invariant term $h_2(\cdot)$, and there are a total of eight to compute the term $P_2(\cdot)u$. Notice that

this number of integrals is only dependent upon the dimension of $f(\cdot)$ and is independent of the number of parameter components or upon any special "controllable" form of $Y_{z.s.}(t;u)$ (see Section IV.3).

To use a gradient method in computing the optimal control for minimizing either $J_1(u)$ or $J_2(u)$ requires that we start with an initial guess of the optimal control (see, e.g., Luenberger (Ref 20: 285)). We choose the initial guess suggested in Section VI.1 as follows:

$$u_o(t) = -h_i(t), \qquad i = 1 \text{ or } 2; \tag{329}$$

for, this is the control which is the unique minimum of the approximate cost functional

$$\int_0^1 (u(t) + h_i(t))u(t)dt \tag{330}$$

Then we may use, for example, a standard steepest descent algorithm (e.g., Luenberger (Ref 20: 285)) to give the sequence of optimal control estimates

$$u_{j+1}(t) = u_j(t) + k_j r_j(t) \tag{331}$$

where $r_j(\cdot)$ is one-half the negative gradient

$$r_j(t) = -(h(t) + P(t)u_j) \tag{332}$$

and

$$k_j = \frac{\int_0^1 r_j(t)r_j(t)dt}{\int_0^1 r_j(t)P(t)r_j dt} \tag{333}$$

The sequence is stopped when the norm of $r_j(\cdot)$ is sufficiently small or when there is no longer any appreciable decrease in the cost functional.

Notice that for each step of iterations (except the first) we must use the subroutine P(t)u just one time since we may write $P(t)u_{j+1}$ as

$$P(t)u_{j+1} = P(t)u_j + k_j P(t)r_j \tag{334}$$

and both $P(t)u_j$ and $P(t)r_j$ have been determined from the previous u_j step. Thus, in this example, for each new iteration of the control there would be four quadrature integrals required if the cost functional is $J_1(u)$ or there would be eight if the cost functional is $J_2(u)$.

The above mentioned steepest descent algorithm was programmed on the CDC 6600 digital computer. Using trapezoidal integration with a step size of .002 on the interval [0, 1], the optimal minimum sensitivity control law was computed. The step size requirement is a function of the accuracy desired, the time constants of the system (the magnitude of the eigenvalues), and the accuracy of the integration scheme selected. The step size of .002 was chosen in order to give reasonable accuracy between the sensitivities computed using the operator form with trapezoidal integration and those computed using the sensitivity system differential equations with a fourth order Runge Kutta integration. However, far more accurate integration techniques then trapezoidal could be utilized and the .002 interval could then, no doubt, be lengthened.

To illustrate the influence of the cost functional upon the speed of convergence, the scalar constants c_1 and c_2 were assigned values of (10, 1) and (100, 10). The optimal control for both the cost functionals $J_1(u)$, equation (318), and $J_2(u)$, equation (319), were computed. The results of the iterations are shown in Tables 1 and 2. Notice that for the set of constants (10, 1) the initial guess $u_o(t) = -h_i(t)$, i=1,2, is

fairly close to the optimal solution, and the optimal control is reached in just three to four steps. For the more heavily weighted constants of (100, 10), the initial guess $u_o(t) = -h_i(t)$, $i = 1, 2$, is further away from the optimal solution and the gradient method takes longer to reach convergence. Notice that for the set of weighting constants selected here, the speed of convergence for both the "terminal" sensitivity problems, $J_1(u)$, or the "trajectory" sensitivity problems, $J_2(u)$, are roughly the same.

Next consider the solution of this example by the Riccati equation technique described in Section VI.2. For this example, the C matrix is the identity matrix and so $Y(t;u) = X(t;u)$. Also, recall from the analysis of this same example in Section IV.3, that the rank of the matrix $[\overline{E}\ \ \overline{G}]$ is $\gamma_1 = 6$; therefore, the eight elements of the sensitivity system, $X(\cdot;u)$, may be calculated from a minimum of six differential equations rather than eight. Indeed, from the results of the example of Section IV.3 we have

$$X(t;u) = \left[\begin{array}{cccccc} & & & I & & \\ \hline 0 & 0 & 2 & 0 & 3 & 1 \\ 0 & 1 & 0 & 2 & -2 & 0 \end{array}\right]_{8\times 6} X_c(t;u)$$

$$\equiv K X_c(t;u) \tag{335}$$

where

$$X_c(t;u) = \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \\ \xi_1^{(1)}(t;u) \\ \xi_2^{(1)}(t;u) \\ \xi_1^{(2)}(t;u) \\ \xi_2^{(2)}(t;u) \end{bmatrix} \tag{336}$$

COMPUTATIONAL RESULTS OF GRADIENT METHOD

Iteration	J_1	$\|\|\nabla J_1\|\|$	J_2	$\|\|\nabla J_2\|\|$
Start	9.378	$.1614 \times 10^{2}$	10.860	$.1801 \times 10^{2}$
1	5.274	.3897	6.322	.4127
2	5.077	$.3736 \times 10^{-1}$	6.121	$.3385 \times 10^{-1}$
3	5.067	$.9017 \times 10^{-3}$	6.112	$.7774 \times 10^{-3}$
4	5.067	$.8646 \times 10^{-4}$	6.112	$.6421 \times 10^{-4}$

Table 1: $c_1 = 10$, $c_2 = 1$

Iteration	J_1	$\|\|\nabla J_1\|\|$	J_2	$\|\|\nabla J_2\|\|$
Start	6049.7	$.1614 \times 10^{6}$	6582.1	1801×10^{6}
1	727.8	$.6552 \times 10^{4}$	708.9	$.6911 \times 10^{4}$
2	117.7	$.2122 \times 10^{4}$	115.5	$.1776 \times 10^{4}$
3	47.73	$.8612 \times 10^{2}$	57.75	$.6831 \times 10^{2}$
4	39.71	$.2790 \times 10^{2}$	51.78	$.1771 \times 10^{2}$
5	38.79	$.1135 \times 10$	51.21	.7577
6	38.68	.3719	51.13	.3056
7	38.67	$.1716 \times 10^{-1}$	51.13	$.7329 \times 10^{-1}$

Table 2: $c_1 = 100$, $c_2 = 10$

Now in terms of $X_c(t;u)$ the cost functionals $J_1(u)$ and $J_2(u)$ may be written as

$$J_1(u) = \int_0^1 u^2(t)dt + c_1[X_c(1;u)/K^TKX_c(1;u)] \tag{337}$$

$$J_2(u) = J_1(u) + c_2 \int_0^1 [X_c(t;u)/K^TKX_c(t;u)]dt \tag{338}$$

The six elements of the "state" $X_c(t;u)$ are then determined from,

$$\dot{X}_c(t;u) = \bar{A}_c X_c(t;u) + \bar{B}_c u(t) \tag{339}$$

$$X_c(0;u) = \bar{d}_c \qquad t \in [0, 1] \tag{340}$$

where

$$\bar{A}_c = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -2 & -3 & 0' & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -2 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 & -3 \end{bmatrix} \quad \bar{B}_c = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \bar{d}_c = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{341}$$

Then the optimal control, $u_2^*(\cdot) \in L_2(0, 1;R)$, which minimizes $J_2(u)$ is uniquely specified by the equation

$$u_2^*(t) = -\bar{B}^T \bar{R}_2(t) X_c^*(t;u_2^*) \qquad t \in [0, 1] \tag{342}$$

where $\bar{R}_2(t)$ is obtained by integrating backwards in time the 6 x 6 dimensioned matrix Riccati differential equation

$$\dot{\bar{R}}_2(t) = -\bar{R}_2(t)\bar{A}_c - \bar{A}_c^T \bar{R}_2(t) + \bar{R}_2(t)\bar{B}_c\bar{B}_c^T \bar{R}_2(t) - c_2K^TK \tag{343}$$

with terminal boundary condition

$$\bar{R}_2(1) = c_1 K^T K \tag{344}$$

The optimal "state" trajectory, $X_c^*(\cdot;u_2^*)$ is obtained by integrating forward in time the equation (339) with the control (342). The optimal control which minimizes $J_1(u)$ is obtained in a similar manner except the differential equation for $\bar{R}_1(t)$, equation (343), no longer has the last term, $c_2 K^T K$. Notice that in either case there are 21 nonlinear differential equations to determine the symmetric $\bar{R}_1(\cdot)$ matrix plus an additional six linear differential equations to determine the optimal trajectory $X_c^*(\cdot;u_1^*)$. Additionally, six function elements of $\bar{R}_1(\cdot)$ must be stored in order to compute the optimal control from equation (342). Finally, the amount of computations are virtually the same regardless of whether the cost functional is $J_1(u)$ or $J_2(u)$. This is in contrast to the gradient method in which the amount of computation was doubled for the trajectory sensitivity cost functional $J_2(u)$ (eight quadratures per iteration versus four for the terminal sensitivity cost functional $J_1(u)$).

3. SUMMARY

In this section we have examined computational techniques for obtaining the open-loop minimum sensitivity control law for the linear time-invariant system, S_{LC}. Using the matrix-operator form of the parameter sensitivities the gradient approach developed in Section II.5 was applied. This proved to be an efficient method of solution in which the amount of computation (number of quadrature integrals) grew only linearly with the state dimension n and the control dimension r. A simple second order example system showed that the computations were quite straight-forward to apply and that convergence on the digital computer could be quite rapid.

Next, this operator-gradient method of solution was compared with well-known Riccati equation solution methods. By applying sensitivity system reduction techniques (Section IV.3) the amount of computations could be reduced substantially. However, the amount of computation (number of nonlinear differential equations for solution of the matrix Riccati differential equation) still grew in proportion to the square of n and r as opposed to linearly with the n and r for the gradient method of solution. Thus, even with sensitivity system reduction methods, the gradient solution offers a practical alternative to computing the open-loop minimum sensitivity control law.

In the next section we consider our final computational application of the matrix-operator representation of the parameter sensitivities in linear time-invariant ordinary differential equation systems. There we examine optimal design of the sensitivities to enhance parameter identification capability, and once again the matrix-operator form provides a useful alternative to differential Riccati equation solution techniques.

Section VII

SENSITIVITY DESIGN FOR OPTIMAL IDENTIFICATION IN LINEAR TIME-INVARIANT ORDINARY DIFFERENTIAL EQUATION SYSTEMS, S_{LC}

In this section we consider the computation of the initial condition vector and control input which optimizes parameter identification capability from the zero-input and zero-state responses, respectively. The methods presented are based upon the discussion of Section II.8, and like the previous two sections we apply the matrix-operator form for the parameter sensitivities in linear time-invariant ordinary differential equation systems (Section III and Appendix A) in order to obtain efficient computational techniques for solution.

The concept of obtaining an optimal initial condition vector is new; however, the approach for obtaining the optimal input function are based upon the work of Mehra (see Section I.1.e and Section II.8 for discussion). Mehra suggests solving the optimal control input design problem via a Rayleigh-Ritz-Galerkin algorithm, and that is the method which we employ here for our operator approach. Therefore, like Section V the contribution here is to demonstrate how the matrix-operator form of the parameter sensitivities can provide a useful computational alternative to previously employed two-point boundary value/Riccati equation methods of solution (e.g., Stepner and Mehra (Ref 120) and Gupta and Hall (Ref 108)). The computational utility of the matrix-operator form is further enhanced through the use of Walsh functions as the orthogonal set of basis functions in the Rayleigh-Ritz-Galerkin algorithm. The special properties of the Walsh functions allow us to substantially reduce the number of integrals required for solution. This method of reduction with the Walsh

functions should have useful applications to other aspects of linear system simulation and control design.

Again the simple second order linear system examined in Sections IV-VI will be used here to illustrate the computational implications of the operator algorithms.

1. OPTIMIZATION OF THE ZERO-INPUT SENSITIVITIES

For the system S_{LC} the zero-input cost functional (see Section II.8.b) takes the form

$$J_{K_{z.i.}}(d) = \sum_{i,j=1}^{p} k_{ij} \int_0^{t_f} [v_{z.i.}^{(i)}(t)/Q(t)v_{z.i.}^{(j)}(t)]dt$$

$$= \sum_{i,j=1}^{p} k_{ij} \int_0^{t_f} [d/(C\Phi(t))_{(i)}^T Q(t)(C\Phi(t))_{(j)} d]dt \qquad (345)$$

The m x m matrix, Q(t), is assumed to be symmetric, positive definite, and uniformly bounded for $t \in [0,t_f]$. It may be assumed that $Q^{-1}(t)$ is the covariance matrix of a Gaussian white noise measurement process (see Section II.6.d).

Using Theorem 4 and defining the n x n positive, symmetric matrix

$$A_{K_{z.i.}} \equiv \sum_{\kappa=1}^{2n} \sum_{\ell=1}^{2n} \{ \sum_{i,j=1}^{p} k_{ij} (CA^{\kappa-1})_{(i)}^T \int_0^{t_f} a_\kappa(t)Q(t)a_\ell(t)dt (CA^{\ell-1})_{(j)} \} \qquad (346)$$

we may write $J_{K_{z.i.}}(d)$ as

$$J_{K_{z.i.}}(d) = [d/A_{K_{z.i.}} d] \qquad (347)$$

Computing the symmetric matrix $A_{K_{z.i.}}$ involves merely matrix manipulations and solving the quadrature integrals

$$\int_0^{t_f} a_\kappa(t)Q(t)a_\ell(t)dt \tag{348}$$

$\kappa = 1, 2, \ldots 2n$

$\ell = 2n, 2n-1, \ldots \kappa$

Also, after solution of these integrals one time, iterative adjustment to the weighting constants, k_{ij}, i,j = 1, 2, ...p, becomes a simple process since these integrals do not change. Therefore, the constants k_{ij} may readily be adjusted to satisfy various optimization criterion (see Section II.8.a).

According to the development of Section II.8.b, we place the auxiliary constraint on the initial condition vector

$$[d/d] \leq 1 \tag{349}$$

We may interpret this constraint as providing the unit vector initial "direction" of maximum information. Then the $d^* \in R^n$ which maximizes $J_{K_{z.i.}}(d)$ is the eigenvector corresponding to the maximum eigenvalue of the positive, symmetric matrix, $A_{K_{z.i.}}$, and normalized so that

$$[d^*/d^*] = 1. \tag{350}$$

Since $A_{K_{z.i.}}$ is positive and symmetric, rapid computer techniques exist for obtaining the maximal eigenvalue and corresponding eigenvector. (See, e.g., Fadeev and Fadeeva, (Ref 11: 406 - 454).) Also, it may be desirable to have an optimization algorithm which will maximize the functional (345) over any given subspace R^n. For example, the Rayleigh-

Ritz-Galerkin algorithm described in detail in the next section has this particular property, and it is easily modified to solve this finite dimensional, matrix eigenvector problem. Then, when implementing a procedure such as suggested in Section IV.2 for reducing the number of influencing parameter components (see the discussion following Corollary 5.1), one may not only select an initial condition vector which is from the insensitive subspace of one group of parameter components, but he may also select the initial condition vector which is optimum with respect to identifying the remaining parameter components.

In conclusion, the method of Section II.8 is well suited for determining an optimal initial condition vector for the system S_{LC}. Indeed, with very little additional computations, the method may be applied to find the optimal initial condition vector for the time-varying linear system, S_{TV}. Such techniques should be valuable when the system designer has the freedom to select the initial condition vector, and he is attempting to obtain the greatest "information" for identifying the system parameters from the transient response.

Example

Consider the second order homogeneous linear system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \qquad t \in [0, 1] \tag{351}$$

with the single observable output

$$y(t) = [1 \quad 0] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{352}$$

The nominal parameter values are $b_1 = -2$ and $b_2 = -3$; thus the nominal A

matrix is the same as the examples of Sections IV, V, and VI and so we will use some of the previous results to shorten the discussion here.

We may then compute

$$(CA^0)_{(1)} = (CA^0)_{(2)} = (CA)_{(1)} = (CA)_{(2)} = 0 \tag{353}$$

$$(CA^2)_{(1)} = [1 \quad 0] \qquad (CA^2)_{(2)} = [0 \quad 1] \tag{354}$$

$$(CA^3)_{(1)} = [-3 \quad 1] \qquad (CA^3)_{(2)} = [-2 \quad -6] \tag{355}$$

Assuming $Q(t) = 1$ for all $t \in [0,1]$ and that

$$k_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{356}$$

we then obtain

$$A_{K_{z.i.}} = \sum_{\kappa,\ell=1}^{4} \sum_{i,j=1}^{2} (CA^{\kappa-1})^T_{(i)} (CA^{\ell-1})_{(j)} \cdot \int_0^1 a_\kappa(t) a_\ell(t) dt$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \int_0^1 a_3(t) a_3(t) dt + \begin{bmatrix} -6 & -1 \\ -1 & -12 \end{bmatrix} \int_0^1 a_3(t) a_4(t) dt + \begin{bmatrix} 13 & 9 \\ 9 & 37 \end{bmatrix} \int_0^1 a_4(t) a_4(t) dt \tag{357}$$

From the inverted Vandermonde matrix of the example in Section V, equation (283), we may determine the scalar functions $a_3(\cdot)$ and $a_4(\cdot)$ to be

$$a_3(t) = -9e^{-t} + 5te^{-t} + 9e^{-2t} + 4te^{-2t} \tag{358}$$

$$a_4(t) = -2e^{-t} + te^{-t} + 2e^{-2t} + te^{-2t} \tag{359}$$

Then

$$\int_0^1 a_3(t)a_3(t) \simeq .022 \tag{360}$$

$$\int_0^1 a_3(t)a_4(t) \simeq .0026 \tag{361}$$

$$\int_0^1 a_4(t)a_4(t) \simeq .0003. \tag{362}$$

Using these values gives the approximate value

$$A_{K_{z.i.}} \simeq \begin{bmatrix} .010 & 0 \\ 0 & .002 \end{bmatrix} \tag{363}$$

Thus $q_M = .01$ is the largest eigenvalue of $A_{K_{z.i.}}$, and the eigenvector corresponding to this eigenvalue is $\beta_M = [1 \quad 0]^T$. Hence the optimal initial condition vector for identifying the parameters in the system (351) - (352) is one selected in the direction β_M.

2. OPTIMIZATION OF THE ZERO-STATE SENSITIVITIES

For simplicity, in this section we will assume a single control input (that is, $r = 1$) and that the weighting matrix Q is time-invariant. The extensions to the more general cases are easily made. Then for the system S_{LC}, the zero-state input design cost functional (see Section II.8.c)

takes the form

$$J_{K_{z.s.}}(u) = \sum_{i,j=1}^{p} k_{ij} \int_0^{t_f} [v_{z.s.}^{(i)}(t;u)/Qv_{z.s.}^{(j)}(t;u)]dt \tag{364}$$

Using the results of Section III and Appendix A, the zero-state sensitivities may be computed from the relations

$$v_{z.s.}^{(i)}(t;u) = H^{(i)} \int_0^t f(t-s)u(s)ds \tag{365}$$

where $f(\cdot)$ is a 2n-dimensioned vector function defined in Appendix A and which has elements of the form $t^j e^{q_k t}$ for real eigenvalue of A, q_k; and $H^{(i)}$, i = 1, 2, ...p, are m x 2n dimensioned matrices computed as described in Appendix A. Using this relationship (365) and interchanging the order of integration, we may put the cost functional (364) into the form

$$J_{K_{z.s.}}(u) = \int_0^{t_f} [u(t)/A_{K_{z.s.}}(t)u]dt \tag{366}$$

where we define the positive self-adjoint operator

$$A_{K_{z.s.}}(t)u = \sum_{i,j=1}^{p} k_{ij} \int_t^{t_f} f^T(s-t)H^{(i)^T} QH^{(j)} \cdot \left(\int_0^s f(s-\zeta)u(\zeta)d\zeta\right)ds \tag{367}$$

Assume that the control must satisfy the "energy" constraint[1]

[1]See footnote of Section II.8.c, page 51.

$$\int_0^{t_f} [u(t)/u(t)]dt \leq 1. \tag{368}$$

Then this problem is a complete parallel to that of Section VII.1. Therefore, the optimal input function which maximizes the cost functional (366) subject to the constraint (368) is the eigenfunction of the positive, self-adjoint, compact operator, $A_{K_{z.s.}}$, which corresponds to its maximal eigenvalue.

This conclusion is a restatement of the results of Mehra (Ref 112), and the motivation for our approach, both in this section and the previous one, is based entirely upon that work. However, the contribution which we wish to make here is to use the new matrix-operator representation of the parameter sensitivities, equation (365), to form an efficient computational algorithm for computing the optimal input function. This algorithm is developed in the next subsection, and then it is compared with previously employed Riccati equation techniques in the Subsection VII.2.b.

a. Rayleigh-Ritz-Galerkin Algorithm for Computing the Maximal Eigenfunction

Mehra (Ref 112) suggests solving the optimal input-maximal eigenfunction problem by means of a Rayleigh-Ritz-Galerkin method of solution; however, he does not present the computational details for this method. The Rayleigh-Ritz-Galerkin method is the solution technique which we elect to adopt using the matrix-operator form of the parameter sensitivities, equation (365). This method has the desirable property that one may select "basis" functions which are easy to implement once the optimal control function is determined as linear combinations basis functions. For example, either sinusoids or Walsh functions (Ref 126) might be utilized as the orthogonal set of basis functions and a linear combination

of either of these might be easy to implement in a physical control system. On the other hand, if the optimal input, u^*, is determined as an arbitrary element of $L_2(0,t_f;R)$, it may be difficult to implement without approximation. Also in using the Rayleigh-Ritz-Galerkin method, the system designer has the freedom to select any level of complexity (and accuracy) of the optimal input by choosing the number of basis functions over which he optimizes.

The details of the Rayleigh-Ritz-Galerkin method of solution are now stated. Notice that in using this technique, we utilize equation (364) rather than (366), as this form proves to be less sensitive to numerical errors, has slightly less total number of integrals, and may easily be extended to time-varying linear systems by computation of the parameter sensitivities from the "sensitivity system".

Rayleigh-Ritz-Galerkin Algorithm

1) Given the nominal A matrix, compute its distinct eigenvalues q_k, $k = 1, 2, \ldots\rho$, and the multiplicities, n_k, of each eigenvalue in the characteristic polynomial of A.

2) Using the eigenvalues of A, form and invert the 2n x 2n Vandermonde matrix Λ as defined in Appendix A. Using the elements of Λ^{-1} and the quantities A, B, C, $A_{(i)}$, $B_{(i)}$, and $C_{(i)}$, $i = 1, 2, \ldots p$, compute the p m x 2n dimensioned matrices[1]

$$H^{(i)} = [(CB)_{(i)} \quad (CAB)_{(i)} \quad \ldots \quad (CA^{2n-1}B)_{(i)}]\Lambda^{-1} \tag{369}$$

[1]Note that this equation holds because r = 1; if r were greater than one (more than one control input) then a similar equation would apply to each column vector of B.

3) For the set of weighting constants $k_{i,j}$, i,j = 1, 2, ...p, and the m x m positive-definite symmetric matrix Q, compute the 2n x 2n dimensioned matrix

$$Q_H = \sum_{i,j=1}^{p} k_{i,j} H^{(i)^T} Q H^{(j)} \tag{370}$$

4) For an arbitrary input function, $u \in L_2(0,t_f;R)$, form a subroutine to compute the 2n-dimensioned vector functions of $t \in [0,t_f]$

$$\phi(t;u) \equiv \int_0^t f(t-s)u(s)ds \tag{371}$$

where the 2n-dimensioned vector functions are defined in Appendix A and have elements of the form $t^j e^{q_k t}$ for real eigenvalue q_k of A (or $t^j e^{\sigma_k t}\cos\omega_k t$ and $t^j e^{\sigma_k t}\sin\omega_k t$ for complex eigenvalues $q_k = \sigma_k \pm \omega_k\sqrt{-1}$). It is important to point out that the indicated 2n convolution integrals of equation (371) may be computed from 2n quadratures through the relations such as

$$(t-s)e^{q_k(t-s)} = te^{q_k t}e^{-q_k s} - e^{q_k t}se^{-q_k s} \tag{372}$$

Comment

At this point we are prepared to begin the actual Rayleigh-Ritz-Galerkin optimization. It is assumed that a set of n_b basis functions, $u_j(t)$, j = 1, 2, ...n_b, $t \in [0,t_f]$ are prespecified. For computational simplicity it is assumed that these basis functions are selected mutually orthogonal; that is, it is assumed that

$$\int_0^{t_f} u_j(t)u_k(t)dt = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{373}$$

5) Compute and store the $n_b \times n_b$ real symmetric matrix, V_Q, with κ^{th} row and ℓ^{th} column element given by

$$V_Q(\kappa,\ell) = \sum_{i,j=1}^{p} k_{i,j} \int_0^{t_f} [v^{(i)}(t;u_\kappa)/Qv^{(j)}(t;u_\ell)]dt$$

$$= \int_0^{t_f} [\phi(t;u_\kappa)/Q_H\phi(t;u_\ell)]dt \qquad (374)$$

6) Compute the maximal eigenvalue, q_M, and its associated normalized eigenvector β_M for the positive symmetric matrix V_Q. Letting $\beta_M(j)$ denote the j^{th} component of β_M, the optimizing control input is then given by

$$u^{\cdot\cdot}(t) = \sum_{j=1}^{n_b} \beta_M(j)u_j(t) \qquad (375)$$

and the optimal value of the cost functional, $J_{K_{z.s.}}(u)$, equation (364), is given by

$$J_{K_{z.s.}}(u^*) = \sum_{j=1}^{n_b} \beta_M(j)V_Q(j,j) \qquad (376)$$

Stop.

Comment 1 Computational Aspects

The preceding algorithm is well suited to digital computer applications. In fact, for a single control input ($r = 1$), the total number of integrals to be computed are, at most:

1) $2nn_b$ to compute the n_b vector functions

$\phi(t;u_j)$, $j = 1, 2, \ldots n_b$, $t \in [0,t_f]$

(see equation (371)). Note that depending upon the basis functions, u_j, these integrals may in fact be computed analytically.

ii) $n_b(n_b + 1)/2$ to compute the symmetric matrix V_Q (see equation (374)).

The remaining computations of the algorithm are straight-forward matrix operations. Also, in the example of Section VII.2.c we will see that the special properties of Walsh basis functions allows us to reduce the number of required integrals in computing the functions $\phi(\cdot;u_j)$, $j = 1, 2, \ldots n_b$, from $2nn_b$ to $2n$. If n_b is large then this is a significant reduction in the amount of computation and storage.

Comment 2 Change of Weighting Constants, k_{ij}

In Section II.8.a we discussed the fact that it is often important to iteratively adjust the weighting constants, k_{ij}, so that the optimal input maximizing the weighted trace cost functional, $J_{K_{z.s.}}(u)$, equation (364), more closely approximates minimization of either the determinant or trace of the inverse information matrix. (See, e.g., Mehra (Ref 114) (Ref 115) or Gupta and Hall (Ref 108).) Therefore, it is important to point out that once the $2nn_b$ integrals

$$\int_0^t f(t - s)u_j(s)ds \qquad (377)$$

$$j = 1, 2, \ldots n_b \qquad t \in [0, t_f]$$

of equation (371) have been computed, they are not affected by adjustment of the k_{ij} constants. Therefore, after the first iteration the amount of computations are significantly reduced for subsequent adjustments of the weighting constants, k_{ij}.

Comment 3 Time-Varying Matrices

The algorithm is easily modified if either the output C matrix or input B matrix are time-varying; the computations are essentially the same except that the indicated integrals must include the time-varying elements of these matrices. However, if the system plant matrix, A, is time-varying then the convenient matrix-operator form (equation (365)) no longer applies. In this case Theorem 1 could be utilized, but it is probably computationally more efficient to compute the parameter sensitivities, $v^{(i)}_{z.s.}(t;u_j)$, required in equation (374), from the sensitivity-system differential equations (89) - (91). Indeed, this approach could also be applied for the time-invariant system S_{LC}, and controllability properties of the sensitivity system (see Section IV.3) could then be used to reduce the required number of differential equations.

b. Solution via the Maximum Principle

As previously stated the theory of the input design problem discussed in Section VII.2 was developed by Mehra (Ref 112). In that paper he reformulates the operator/eigenfunction problem into a differential equation two-point boundary value problem through the use of the "sensitivity system" and the Maximum Principle (Ref 22). For the control constraint $[u,u] \leq 1$, he shows that the optimal input satisfies the two-point boundary value problem

$$\frac{d}{dt}\begin{bmatrix} X(t;u) \\ \lambda(t) \end{bmatrix} = \begin{bmatrix} \bar{A} & -\alpha\bar{B}\bar{B}^T \\ \bar{C}^T\bar{Q}\bar{C} & -\bar{A}^T \end{bmatrix}\begin{bmatrix} X(t;u) \\ \lambda(t) \end{bmatrix} \equiv H\begin{bmatrix} X(t;u) \\ \lambda(t) \end{bmatrix} \tag{378}$$

$$X(0;u) = 0 \qquad \lambda(t_f) = 0 \tag{379}$$

$$u^*(t) = -\alpha^* \bar{B}^T \lambda(t) \tag{380}$$

where $\lambda(\cdot)$ is an $n(p+1)$ dimensioned Lagrange multiplier function; $\bar{Q}$ is given by

$$\bar{Q} = \begin{bmatrix} Q & 0 & \cdots & 0 \\ 0 & Q & & \\ \vdots & & & \vdots \\ 0 & \cdots & & Q \end{bmatrix}_{m(p+1)xm(p+1)} \tag{381}$$

and $X(t;u)$, $\bar{A}$, $\bar{B}$, and $\bar{C}$ are defined by expressions (92) - (94) of Section III. The smallest value of constant α for a nontrivial solution to (378) - (380) gives the optimal control input computed from equation (380). Mehra also shows that the optimum α^* is inversely proportional to the optimum value of the cost functional (trace of the weighted local information matrix).

Mehra (Ref 112) suggests solving this two-point boundary value problem by means or a Riccati equation method, and in Reference 120 Stepner and Mehra present a detailed computational algorithm using this technique. However, like the Riccati equation method used in Section VI.3 for minimum sensitivity control design, this approach requires the precise integration of the nonlinear matrix Riccati equation of order $n(p+1) \times n(p+1)$. This, coupled with other computational problems, makes the Stepner and Mehra algorithm somewhat less than satisfactory, (Ref 108). In more recent work Gupta and Hall (Ref 108) have used an eigenvalue-eigenvector decomposition of the Hamiltonian matrix, H, to solve the two-point boundary value

problem, and thereby they avoid integration of the nonlinear Riccati differential equations. This technique, coupled with taking the controllability properties into account to eliminate redundant calculations in the sensitivity system (see Section IV.3) makes the total computational requirements of the Gupta and Hall algorithm roughly equivalent to those of the algorithm of Section VII.2.a. However, there are some other important differences between the two methods of solution.

The first difference is that instead of specifying the time interval of observation $[0, t_f]$ (as is the case in our algorithm), the two-point boundary value method specifies the constant value α^*. In essence, this specifies the optimal value of the cost functional (trace of the weighted information matrix), and the resulting optimal control and time interval of observation $[0, t_f^*]$ are computed from this constant value α^*. If this is not the desired time interval, then α^* must be adjusted and the problem recomputed. In certain cases it may be convenient to specify the level of the "information" required from an experiment and the time interval of observation fixed according to this level; however, in other cases the time interval of observation must be fixed and one must use whatever information is obtainable on this interval. In the first instance the algorithm of Gupta and Hall might be the more direct while in the latter case the algorithm of Section VI.2.a might be more straight-forward to utilize; however, in either case the dependent parameter may be adjusted by suitable iteration of the independent parameter.

The second difference is simply one of implementing an exact optimal control function as obtained by the Gupta and Hall method versus the implementation of control law which is formed by the optimal linear combination of preselected basis functions. It was mentioned previously

that the basis functions used in the Rayleigh-Ritz-Galerkin algorithm might be selected with some physical constraints or other considerations in mind, and this might facilitate the actual implementation of the control law. Again this difference is not of major importance because the "exact" optimal control may be approximated on the one hand, and any level of "optimality" may be achieved, on the other hand, simply by increasing the number of basis functions over which the optimization is performed.

The final major difference is in the amount of computations required to recompute a new optimal input once the weighting constants k_{ij} are adjusted. In the Gupta and Hall algorithm the amount of computation remains the same for each new set of weightings, whereas the amount of computation via the algorithm of Section VII.2.a is significantly reduced once the first optimal value has been obtained (see Comment 2 of Section VII.2.a). Since it is often quite important to make such adjustments to the weighting constants (see, e.g., Mehra (Ref 114) (Ref 115) or Gupta and Hall (Ref 108)), this is a significant difference between the two algorithms.

c. Example

To illustrate the use of the Rayleigh-Ritz-Galerkin algorithm of Section VII.2.a, we once again consider the second order time-invariant system:

$$\begin{bmatrix} \dot{x}_1(t;u) \\ \dot{x}_2(t;u) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \end{bmatrix} + \begin{bmatrix} 0 \\ b_3 \end{bmatrix} u(t) \tag{382}$$

$$\begin{bmatrix} y_1(t;u) \\ y_2(t;u) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t;u) \\ x_2(t;u) \end{bmatrix} \qquad t \in [0, 8] \tag{383}$$

with zero initial conditions. Also, in order to illustrate one time the computational features associated with complex eigenvalues, we will here assume that the nominal parameter vector is $b_o = [-3 \quad -2 \quad 1]^T$. Therefore, the nominal eigenvalues are

$$q_1 = \sigma + \omega\sqrt{-1} = -1 + \sqrt{2}\ \sqrt{-1} \tag{384}$$

$$q_2 = \sigma - \omega\sqrt{-1} = -1 - \sqrt{2}\ \sqrt{-1} \tag{385}$$

Then the generalized Vandermonde matrix, Λ, described in Appendix A becomes

$$\Lambda = \begin{bmatrix} 1 & \sigma & \sigma^2 - \omega^2 & \sigma^3 - 3\sigma\omega^2 \\ 0 & \omega & 2\sigma\omega & 3\sigma^2\omega - \omega^3 \\ 0 & 1 & 2\sigma & 3(\sigma^2 - \omega^2) \\ 0 & 0 & 2\omega & 6\sigma\omega \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & -1 & 5 \\ 0 & \sqrt{2} & -2\sqrt{2} & \sqrt{2} \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 2\sqrt{2} & -6\sqrt{2} \end{bmatrix} \tag{386}$$

and the vector function $f(t)$ is

$$f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{bmatrix} = \begin{bmatrix} e^{-t}\cos\sqrt{2}\,t \\ e^{-t}\sin\sqrt{2}\,t \\ te^{-t}\cos\sqrt{2}\,t \\ te^{-t}\sin\sqrt{2}\,t \end{bmatrix} \tag{387}$$

Since there is a single control input, the matrices $H^{(i)}$, i = 1, 2, 3, are computed from the expression

$$H^{(i)} = [(CB)_{(i)} \quad (CAB)_{(i)} \quad (CA^2B)_{(i)} \quad (CA^3B)_{(i)}] \Lambda^{-1} \tag{388}$$

Then assuming that Q is the identity matrix and that

$$k_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{389}$$

The matrix Q_H becomes

$$Q_H = \sum_{i,j=1}^{3} k_{ij} H^{(i)^T} Q H^{(j)}$$

$$= \sum_{i=1}^{3} H^{(i)^T} H^{(i)} \tag{390}$$

This matrix was computed on the digital computer to be

$$Q_H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & .8125 & .1326 & 0 \\ 0 & .1326 & .1875 & .1768 \\ 0 & 0 & .1768 & .25 \end{bmatrix} \tag{391}$$

With these preliminaries we are ready to begin the Rayleigh-Ritz-Galerkin optimization. To do so requires that we select a set of orthogonal basis functions, $\{u_j(\cdot) \in L_2(0,8;R);\ j = 1, 2, \ldots n_b\}$, over which we desire to optimize. For example, sinusoids could be selected and such functions have found frequent application in parameter identification problems. However, it will be particularly illuminating to select an orthogonal set of Walsh functions (Ref 126)) as our optimization set. We will see that

the special properties of the Walsh functions enables us to reduce the overall computational requirements substantially. Indeed, the method by which we reduce the computations is useful in its own right, and should contribute to already existing methods for utilizing Walsh functions in system simulation and control function design. (See, e.g., Chen and Hsiao (Ref 123) (Ref 124) and Corrington (Ref 125) where Walsh functions are used to approximate the solution of differential equation representations of dynamic systems.)

Therefore, we choose as our basis set the first sixteen Walsh functions. These may be defined by the following 16 x 16 dimensioned symmetric matrix:

$$W_a = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1
\end{bmatrix} \tag{392}$$

Each of the column vectors of this 16 x 16 matrix corresponds to one Walsh function, and each row corresponds to the function values on the corresponding one-half second interval of time.

Now let $w_j = [w_j(1) \quad w_j(2) \ \ldots \ w_j(16)]^T$ be the j^{th} column vector of W_a, and define the 16 x 1 dimensioned vectors

$$\theta_k = [\int_0^{1/2} f_k(-s)ds \quad \int_{1/2}^{1} f_k(-s)ds \ \ldots \ \int_{15/2}^{8} f_k(-s)ds]^T.$$

$$k = 1, 2, 3, 4 \tag{393}$$

where $f_k(\cdot)$, $k = 1, 2, 3, 4$, are the four elements of $f(\cdot)$, equation (387). Notice that each of the integrals in θ_k may be computed analytically. In order to compute the optimal solution via the Rayleigh-Ritz-Galerkin method of solution, we must compute the vector functions of $t \in [0, 8]$

$$\phi(t;u_j) = \int_0^t f(t - s)u_j(s)ds \tag{394}$$

where $u_j(\cdot) \in L_2(0, 8; R)$, $j = 1, 2, \ldots 16$, are the sixteen Walsh basis functions. Using the relation

$$f(t-s) = e^{-t} \begin{bmatrix} \cos\sqrt{2}\,t & \sin\sqrt{2}\,t & 0 & 0 \\ \sin\sqrt{2}\,t & -\cos\sqrt{2}\,t & 0 & 0 \\ t\cos\sqrt{2}\,t & t\sin\sqrt{2}\,t & \cos\sqrt{2}\,t & \sin\sqrt{2}\,t \\ t\sin\sqrt{2}\,t & -t\cos\sqrt{2}\,t & \sin\sqrt{2}\,t & -\cos\sqrt{2}\,t \end{bmatrix} \begin{bmatrix} e^s\cos\sqrt{2}\,s \\ e^s\sin\sqrt{2}\,s \\ -s\,e^s\cos\sqrt{2}\,s \\ -s\,e^s\sin\sqrt{2}\,s \end{bmatrix}$$

$$\equiv \Psi(t)f(-s) \tag{395}$$

we may then compute the vector functions $\phi(\cdot;u_j) \in L_2(0, 8; R^4)$, $j = 1, 2, \ldots 16$, from the relations

$$\phi(t;u_j) = \int_0^t f(t-s)u_j(s)ds$$

$$= \Psi(t) \int_0^t f(-s)u_j(s)ds \qquad (396)$$

where the k^{th} element of this integral is computed from

$$\int_0^t f_k(-s)u_j(s)ds = \sum_{\ell=0}^{L} w_j(\ell+1)\theta_k(\ell+1) + w_j(L+1)\int_{L/2}^t f_k(-s)ds \qquad (397)$$

Therefore, to compute all $n_b = 16$ vector functions $\phi(t;u_j)$, $t \in [0, 8]$, it now only requires the $2n = 4$ integrals over the interval $[0, 8]$:

$$\int_{L/2}^t f_k(-s)ds \qquad t \in [L/2, L/2 + 1/2]$$

$$L = 0, 1, \ldots 15$$

$$k = 1, 2, 3, 4 \qquad (398)$$

That is, because we are dealing with Walsh functions (functions which are either +1 or -1 over finite intervals), and because we are able to write $f(t-s)$ in the factored form $\Psi(t)f(-s)$, the integrals in computing the functions $\phi(t;u_j)$, $j = 1, 2, \ldots 16$, are independent of specific u_j. Furthermore, the integrals (398) may all be computed analytically. Thus while there would normally be $2nn_b = 64$ integrals required to compute the 16 individual vector functions $\phi(t;u_j)$, there are now only $2n = 4$ integrals required and these may all be computed analytically. This is quite a substantial reduction in computation and storage which is associated with the use of Walsh functions in the Rayleigh-Ritz-Galerkin optimal input design algorithm. We also comment that this same operator technique could be applied to other simulation and control design problems which are associated

with the use of Walsh functions with linear time-invariant systems. (See, e.g., Chen and Hsiao (Ref 123) (Ref 124) and Corrington (Ref 125)).

To complete the Rayleigh-Ritz-Galerkin algorithm requires that we compute the symmetric 16 x 16 dimensioned matrix, V_Q, equation (374), with i^{th} row and j^{th} column element

$$V_Q(i,j) = \int_0^8 [\phi(t;u_i)/Q_H\phi(t;u_j)]dt \tag{399}$$

This requires $n_b(n_b + 1)/2 = 136$ integrals which cannot be reduced. Once computing this matrix, we then determine its maximal eigenvalue and corresponding normalized eigenvector, $\beta_M \in R^{16}$. The optimal input is then computed to be

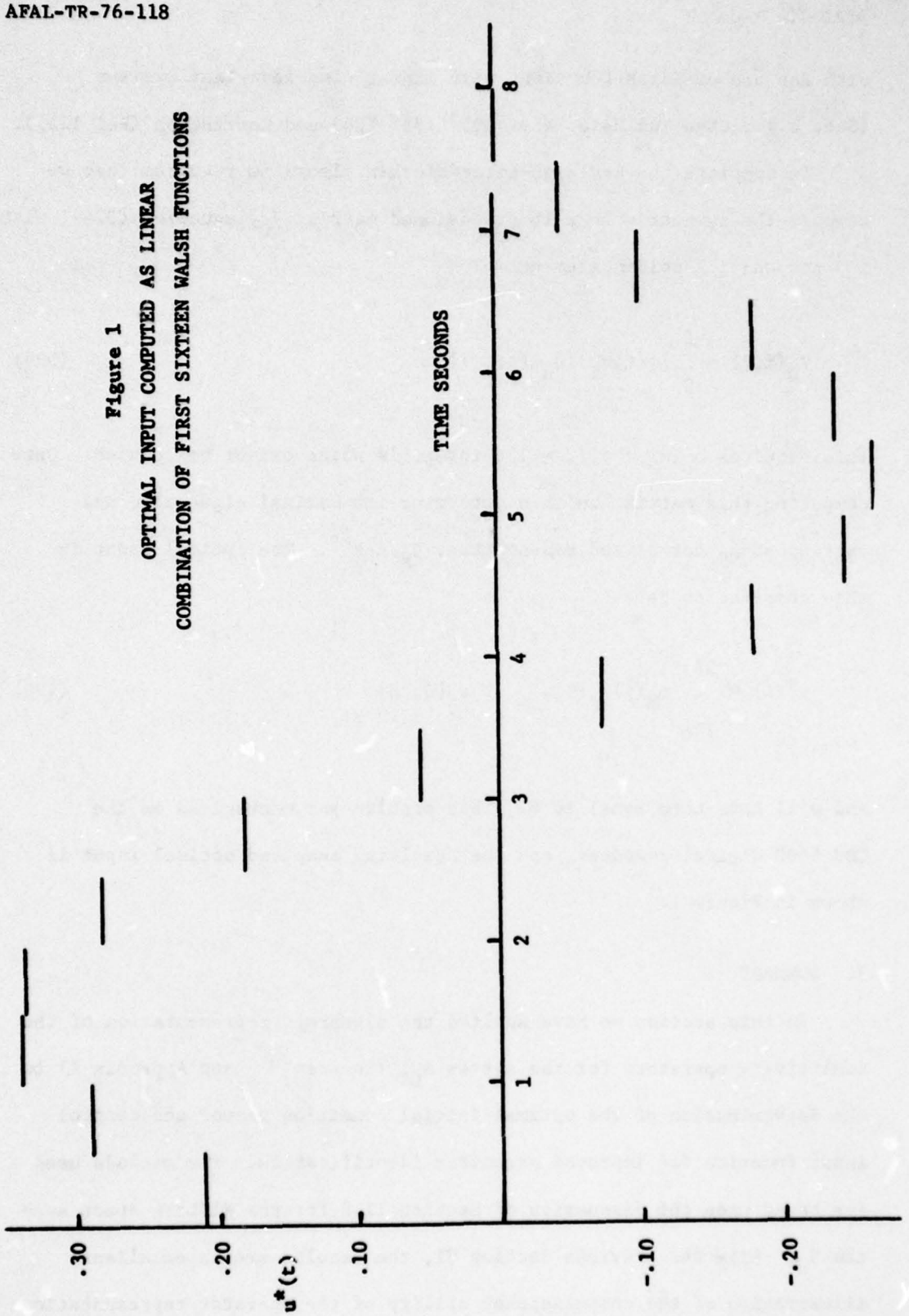
$$u^*(t) = \sum_{j=1}^{16} \beta_M(j)u_j(t), \qquad t \in [0, 8] \tag{400}$$

and will have norm equal to 8. This problem was mechanized on the CDC 6600 digital computer, and the resulting computed optimal input is shown in Figure 1.

3. SUMMARY

In this section we have applied the algebraic representation of the sensitivity operators for the system S_{LC} (Theorem 4 and Appendix A) to the determination of the optimal initial condition vector and control input function for improved parameter identification. The methods used are based upon the discussion of Section II.8 for the Hilbert space system S_H. Like the previous Section VI, the results are an excellent illustration of the computational utility of the operator representation

Figure 1
OPTIMAL INPUT COMPUTED AS LINEAR
COMBINATION OF FIRST SIXTEEN WALSH FUNCTIONS



and approach. For optimal control input design the Rayleigh-Ritz-Galerkin method is the selected computational procedure, and this is shown to be a viable alternative to previously employed Riccati equation techniques. Finally, if Walsh functions are selected as the set of basis functions for the Rayleigh-Ritz-Galerkin optimization, then considerable computational reductions are derived. These reduction methods should also have general application to the use of Walsh functions for system simulation and control design in linear time-invariant systems.

Section VIII

SUMMARY AND RECOMMENDATIONS FOR FUTURE WORK

In this section we review the major objectives and contributions of this research, and then discuss some potential areas of extension and future research.

1. OBJECTIVES AND CONTRIBUTIONS

There were two major objectives in this research and presentation. The first was to demonstrate the useful way in which an operator approach could help understand and investigate system properties of parameter sensitivity in linear dynamic systems. The sensitivities themselves are linear operators (Fréchet derivatives) and so this is a natural way in which to treat them. In Section II the tools of functional analysis were applied to investigate many of their system properties for a general Hilbert space linear system. Of particular prominence in this analysis was the adjoint operator of the sensitivity operators; the adjoint operator played a major role in every aspect of sensitivity theory which was discussed in Section II. Section II, then, laid the ground work for this operator approach to treating the parameter sensitivities and established important methodology for application to linear ordinary differential equation systems in Sections III - VII.

The second objective was to use this operator approach in the development of practical computational algorithms for application to simulation, analysis, identification, and control of linear ordinary differential equation systems. It is felt that the new matrix-operator description of the parameter sensitivities in linear time-invariant sys-

tems (Theorems 2 - 4) developed in Section III contributed to realizing this second objective. This form provided extensions to existing structural information concerning the sensitivity system (Guardabassi, et al (Ref 51) (Ref 52)) and to obtain a new and completely general method for computing the eigenvalue sensitivities (Section III.2).

Further geometrical properties of this matrix-operator form of the parameter sensitivities were then studied in Section IV. Some previously known results concerning insensitivity (Theorems 5 and 6) and sensitivity controllability (Theorem 7) were derived in a more direct fashion by using this form. Discussing the relationship between insensitivity and identifiability, a systematic way of designing identification experiments which would reduce the number of "influencing" parameter components was discussed (Corollary 5.1). A new result concerning necessary and sufficient conditions for zero terminal sensitivity was derived (Corollary 7.1), and the relationship between insensitivity and sensitivity controllability was discussed. Next, new local and structural conditions for zero-input and zero-state identifiability were obtained (Theorems 8 and 9). These conditions were completely general, were only dependent upon the structure of the system matrices, and required no integrals to compute. Finally, the relationship between the control input and steady-state identifiability was discussed and given new insight from the matrix-operator form of the parameter sensitivities.

The computational applications of this matrix-operator form were next considered. In Section V this form was used in a quasilinearization algorithm for parameter identification in linear time-invariant systems. Each iteration of the algorithm required that the new eigenvalues of the A matrix be computed and, for state dimension n and control dimension r,

there were but 2nr quadrature integrals required to compute the complete set of parameter sensitivities. Based upon equation (157) of Section III it was shown how the eigenvalue sensitivities could be obtained directly from the matrix description of the output parameter sensitivities and thes eigenvalue sensitivities could then be used to help iteratively compute the new eigenvalues. Also, based upon methods presented in Section IV.3, linear dependency of row vectors in the matrices of the operator description could be used to reduce the number of integrals required to compute the local information matrix and local gradient function for each iteration of the quasilinearization algorithm. These reduction techniques were similar in approach to methods applied by Gupta and Mehra (Ref 84) for reducing computations in differential equation descriptions of the parameter sensitivity system.

Next the matrix-operator form of the parameter sensitivities in linear time-invariant ordinary differential equation systems was used as the basis of a gradient method for computing the optimal open-loop minimum sensitivity control law (Section VI). This gradient method of solution was developed in Section II.5. By using the matrix-operator description of the parameter sensitivities obtained in Section III, the computational requirements of this method proved to be quite efficient and the computations increased only linearly with state dimension p. This was in contrast to previously employed Riccati equation methods of solution (e.g., Kahne (Ref 70)) in which, even with the use of minimal order sensitivity models, the amount of computation increases with the square of n and square of r.

The final application of this matrix operator form was in the computation of the optimal initial condition vector and optimal control

input to improve parameter identification from the zero-input and zero-state responses, respectively. The theory of the methods was based upon results of Mehra (Ref 112) for optimal input design from the zero-state response, and so the main contribution here was to show the computational utility of the matrix-operator form of the parameter sensitivities. For the optimal input design problem, a Rayleigh-Ritz-Galerkin optimization method was selected. Using the matrix-operator form of the parameter sensitivities, this optimization yielded a highly efficient and computationally straight-forward method of solution. Additionally, if Walsh functions were selected as the set of basis functions for the Rayleigh-Ritz-Galerkin procedure, then the computational requirements could be reduced still further. An example illustrating the method was presented, and the technique for reducing the computations with the use of Walsh functions should have further applications to system simulation and control design in linear time-invariant ordinary differential equation systems.

In all of these applications to linear time-invariant ordinary differential equation systems in Sections III - VII, the operator methodology was used as an underlying foundation. It is hoped that the methods and results presented here will be computationally useful and will spur new research into this type of an approach for treating parameter sensitivity in linear dynamic systems. Just a very few of the unanswered questions and potential avenues for further research are touched upon in the following subsection.

2. EXTENSIONS AND FUTURE RESEARCH

Although the sensitivity-related system properties developed in Section II were done for a very general Hilbert space linear system, they were applied only to linear ordinary differential equation systems

in Sections III - VII. Therefore, it would be quite desirable to apply these methods to linear hereditary and partial differential equation systems as well. (See, for example, (Ref 1).) Of particular significance in this regard might be the design of a control input to optimize parameter identification capability in such systems; the Rayleigh-Ritz-Galerkin algorithm could readily be applied as a solution method for such systems. The gradient approach to minimum sensitivity control design might also be easily applied, particularly for the minimum "terminal" sensitivity problem. Indeed, in Section VI.1 the point was made that the minimum terminal sensitivity control problem inherently involves no convolution integrals. This is particularly important for such general systems because, like the case for a time-varying plant matrix, such convolutions probably cannot be transformed into the far more desirable quadrature integral form as was done with the time-invariant system S_{LC}.

The next major area of extension would be to consider minimum sensitivity control design when the control law is closed-loop or possesses feedback. This subject has been treated extensively through "sensitivity system" differential equation approaches (see, for example, (Ref 69) (Ref 72) or (Ref 75)), and the operator philosophy and the algebraic representation of the sensitivity operators for linear time-invariant systems may yield new approaches and attractive computational techniques. We have carefully restricted our attention to "open-loop" control design, as the closed-loop questions are rich in both theoretical and practical considerations and to delve into them in any depth would have expanded our scope considerably.

A topic area similar to closed-loop minimum sensitivity control

design in that it is abundant with theoretical considerations is the theory of dual control first introduced by Fel'dbaum (Ref 12). Underlying the concept of dual control is the fact that the control law which optimizes some particular control objective may not be optimal with respect to estimation of the system characteristics, parameters, or states. Indeed, this fact is emphasized by the conflicting goals of having the sensitivity operators "large" for improved identification but "small" for minimum trajectory perturbations. Since the knowledge of the system parameters and state inherently affects the ability to achieve the control objective, a truly optimal control must be one which takes both the control and estimation objectives into account. The optimal formulation of dual control results in an infinite dimensional problem and even approximate methods of solution present severe computation burdens. (See, e.g., (Ref 30).) However, parameter sensitivity is inherently related to both the identification of parameters and the design of control laws which are "forgiving" to uncertainty in the exact knowledge of the parameter values. Therefore, techniques designed around the use of the matrix-operator representation of the parameter sensitivities in linear time-invariant systems might yield adaptive control laws having approximately optimal dual control characteristics and which are also easy to compute.

Another area which appears to be quite fruitful for further research is a frequency domain description of the parameter sensitivities and an analysis of their system properties in the frequency domain. The operator time-domain approach for treating the sensitivities is closely akin to a frequency domain method, and the algebraic representation which we have obtained for the parameter sensitivities of the time-invariant sys-

tem S_{LC} (Section III) can also be derived in the frequency domain through partial fraction expansion methods. Frequency domain methods have been very effective for the analysis and design of feedback control laws for minimum sensitivity (e.g., Horowitz (Ref 14) and Horowitz and Shaked (Ref 15)). Additionally, Mehra (Ref 114) (Ref 116) has used frequency domain methods quite beneficially in the design of control laws to optimize parameter identification capability. Further work relating the time-domain and frequency-domain approaches would be most beneficial.

The final, and perhaps most important area of extension would be to use the methods and computational techniques of Sections III - VII to solve practical problems in simulation, identification, and control - particularly in large scale systems where the state dimension n and parameter dimension p are large. Really, it is here in such large scale systems where the algorithms and results obtained should have their most beneficial payoff. In Section IV the properties of insensitivity and identifiability were related, and techniques were suggested for the design of parameter identification experiments which might reduce the number of "influencing" parameters in each experiment. Such methods would be particularly significant in large dimension systems and practical examples demonstrating the concepts are most important in establishing the theory. Also, in each of the operator algorithms for identification and control suggested in Sections V - VII, a large number of differential equations by conventional techniques were replaced by a comparatively small number of quadrature integrals. Since obtaining such practical algorithms is a major aim of this research, it is hoped that these results will indeed prove to be useful tools for practical applications in systems theory.

Appendix A

COMPUTATIONAL DETAILS FOR OBTAINING PARAMETER SENSITIVITIES FOR THE SYSTEM S_{LC}

Let the real n x n matrix A have the characteristic polynomial

$$\Delta(q) = \prod_{k=1}^{\rho} (q - q_k)^{n_k} \tag{401}$$

where the eigenvalues $q_1, q_2, \ldots q_\tau$ are real and the $q_k = \sigma_k + \omega_k\sqrt{-1}$ and $q_{k+1} = \sigma_k - \omega_k\sqrt{-1}$ for $k = \tau + 1, \tau + 3, \ldots \rho-1$. Define the generalized Vandermonde matrix, Λ, with row entries

$$\frac{d^j}{dq^j} \left[1 \quad q \quad q^2 \ldots q^{2n-1}\right]_{q=q_k} \tag{402}$$

for $k = 1, 2, \ldots \tau$, $j = 0, 1, \ldots 2n_k-1$ and double row entries

$$\begin{aligned} &\text{Real } \left(\frac{d^j}{dq^j} \left[1 \quad q \quad q^2 \ldots q^{2n-1}\right]_{q=q_k}\right) \\ &\text{Imag } \left(\frac{d^j}{dq^j} \left[1 \quad q \quad q^2 \ldots q^{2n-1}\right]_{q=q_k}\right) \end{aligned} \tag{403}$$

for $k = \tau + 1, \tau + 3, \ldots \rho-1$, $j = 0, 1, \ldots 2n_k-1$. Next define the 2n dimensioned vector functions

$$a(t) \equiv \left[a_1(t) \quad a_2(t) \ldots a_{2n}(t)\right]^T \tag{404}$$

$$f(t) \equiv \begin{bmatrix} e^{q_1 t} \\ te^{q_1 t} \\ \vdots \\ t^{2n_1-1} e^{q_1 t} \\ \vdots \\ e^{\sigma_{\tau+1} t} \cos\omega_{\tau+1} t \\ e^{\sigma_{\tau+1} t} \sin\omega_{\tau+1} t \\ te^{\sigma_{\tau+1} t} \cos\omega_{\tau+1} t \\ te^{\sigma_{\tau+1} t} \sin\omega_{\tau+1} t \\ \vdots \end{bmatrix} \tag{405}$$

Then we may determine the scalar functions, $a_j(\cdot)$, of Theorem 4 from the equation

$$a(t) = \Lambda^{-1} f(t) \tag{406}$$

However, in computational applications we will generally not use Theorem 4 and the scalar functions $a_j(\cdot)$. Rather we will use a "component" representation of the parameter sensitivities, as such a form allows us to directly convert convolution integrals into quadrature integrals. To this end, define the $2n(p + 1)$ m-dimensional vectors

$$F^{(0)}(k) \equiv \sum_{j=1}^{2n} (CA^{j-1}d)\Lambda^{-1}(j,k) \tag{407}$$

$$F^{(i)}(k) \equiv \sum_{j=1}^{2n} (CA^{j-1})_{(i)} \Lambda^{-1}(j,k) \tag{408}$$

and the 2n(p + 1) m x r-dimensioned matrices

$$H^{(0)}(k) \equiv \sum_{j=1}^{2n} (CA^{j-1}B)\Lambda^{-1}(j,k) \tag{409}$$

$$H^{(i)}(k) \equiv \sum_{j=1}^{2n} (CA^{j-1}B)_{(i)} \Lambda^{-1}(j,k) \tag{410}$$

$i = 1, 2, \ldots p, \; k = 1, 2, \ldots 2n$

Note that the partial derivatives may be computed recursively on the digital computer by realizing that

$$(A^{j+1})_{(i)} = A(A^{j})_{(i)} + A_{(i)} A^{j} \tag{411}$$

Next we define the matrices

$$F^{(i)} \equiv [F^{(i)}(1) \quad F^{(i)}(2) \ldots F^{(i)}(2n)]_{mx2n} \tag{412}$$

$$H^{(i)} \equiv [H^{(i)}(1) \quad H^{(i)}(2) \ldots H^{(i)}(2n)]_{mx2nr} \tag{413}$$

$i = 0, 1, \ldots p$

$$\overline{F} \equiv \begin{bmatrix} F^{(0)} \\ F^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ F^{(p)} \end{bmatrix}_{m(p+1)x2n} \qquad \overline{H} \begin{bmatrix} H^{(0)} \\ H^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ H^{(p)} \end{bmatrix}_{m(p+1)x2nr} \tag{414}$$

the augmented vector

$$Y(t;u) = [y^T(t;u) \quad v^{(1)^T}(t;u) \ \ldots \ v^{(p)^T}(t;u)]^T \tag{415}$$

and the special product

$$f(t)*u(t) \equiv \begin{bmatrix} f_1(t)u(t) \\ f_2(t)u(t) \\ \vdots \\ f_{2n}(t)u(t) \end{bmatrix}_{2nrx1} \tag{416}$$

Then in terms of these quantities we may rewrite Theorem 4 in the compact matrix-operator form:

Theorem A.1

For the system S_{LC}

$$Y(t;u) = \bar{F}f(t) + \bar{H}\int_0^t f(t-s)*u(s)ds \tag{417}$$

This form is entirely equivalent to Theorem 4 but is most convenient for digital computer computation because the elements of the vector function f(t) are the individual modal functions, $t^j e^{q_k t}$ (or cosines and sines for those eigenvalues which are complex conjugate conjugate-pairs). This fact allows us to directly transform the convolutions of equation (417) into quadratures through use of relationships such as

$$(t-s)e^{q_k(t-s)} = te^{q_k t}e^{-q_k s} - e^{q_k t}se^{-q_k s} \tag{418}$$

This is an extremely important fact and we cannot overstate the computational significance of this simple relationship.

From Theorem A.1 we see that the entire sensitivity system may be generated with at most 2nr quadrature integrals. This number of integrals may be reduced to a minimum by elimination of column vectors of $\bar{H}$ which are identically zero. In fact, we are assured that the column "matrix"

$$\bar{H}(k) \equiv \begin{bmatrix} H^{(0)}(k) \\ H^{(1)}(k) \\ \vdots \\ H^{(p)}(k) \end{bmatrix}_{m(p+1)xr} \tag{419}$$

will be identically zero if k corresponds to a multiplicity of one of the eigenvalues which is greater than that eigenvalue s multiplicity in the minimal polynomial of '

$$\bar{A} \equiv \begin{bmatrix} A & 0 & \cdots & 0 \\ A_{(1)} & A & & \vdots \\ \vdots & \vdots & & \vdots \\ A_{(p)} & 0 & \cdots & A \end{bmatrix}_{n(p+1)xn(p+1)} \tag{420}$$

In fact, letting $\bar{\mu}$ be the order of the minimal polynomial of $\bar{A}$, we are assured at least $2n - \bar{\mu}$ of these columns, $\bar{H}(k)$, will be identically zero (or approximately so when considering numerical errors on the digital computer). Therefore, the sensitivity Y(t;u) may be computed with only $\bar{\mu}r$ quadratures instead of 2nr quadratures.

Appendix B

SECOND ORDER SENSITIVITY OPERATORS FOR THE LINEAR TIME-INVARIANT SYSTEM S_{LC}

In a parallel, but somewhat abbreviated, manner to Section III we will develop operator expressions for the second-order sensitivities

$$v^{(i,j)}(t) \equiv \left.\frac{\partial^2 y(t;b)}{\partial b_i \partial b_j}\right|_{b=b_o}$$

$$= y_{(i,j)}(t) \tag{421}$$

Using the second order Taylor's series approximation, it is seen that such second order sensitivity operators provide a more accurate prediction of the output perturbation caused by small parameter variations:

$$y(t;b) = y(t;b_o) + \sum_{i=1}^{p} v^{(i)}(t)\Delta b_i + 1/2 \sum_{i=1}^{p}\sum_{j=1}^{p} v^{(i,j)}(t)\Delta b_i \Delta b_j$$

$$+ \varepsilon_2(t) \tag{422}$$

where ε_2 is of third order in Δb. Also, in Section II.6.c the second order sensitivity operators were required for the Newton-Raphson parameter identification algorithm.

If it is assumed that the second partial derivatives of A, B, and C exist and are continuous at b_o, then the second order sensitivities may be obtained from the "sensitivity system" differential equations

AD-A034 057 AIR FORCE AVIONICS LAB WRIGHT-PATTERSON AFB OHIO F/G 12/1
SENSITIVITY OPERATORS AND ASSOCIATED SYSTEM CONCEPTS FOR LINEAR--ETC(U)
JUL 76 J G REID
UNCLASSIFIED AFAL-TR-76-118 NL

3 OF 3
AD-A
034 057

END
DATE
FILMED
2-15-77
NTIS

$$\dot{\xi}^{(i,j)}(t) = A\xi^{(i,j)}(t) + A_{(i)}\xi^{(j)}(t) + A_{(j)}\xi^{(i)}(t) + A_{(i,j)}x(t)$$

$$+ B_{(i,j)}u(t) \tag{423}$$

$$\xi^{(i,j)}(0) = d_{(i,j)} \tag{424}$$

$$v^{(i,j)}(t) = C\xi^{(i,j)}(t) + C_{(i)}v^{(j)}(t) + C_{(j)}v^{(i)}(t) + C_{(i,j)}x(t) \tag{425}$$

The solution of these equations together with the solution of the $n(p + 1)$ differential equations to obtain x and $\xi^{(i)}(t)$, $i = 1, 2, \ldots p$, requires a total of at most $n(p + 1 + p!)$ coupled linear differential equations. "Minimal-order" models of the second order sensitivities have been obtained (Bingulac (Ref 34) Varshney (Ref 40)), but by and large there has been relatively little investigation of the second order sensitivities. Here we show that results parallel to Section III can be used to compute the first and second order sensitivities with at most 3nr quadrature integrals.

As in Section III the first step in obtaining an operator expression for the second order sensitivities is the determination of the second partial derivatives of the state transition matrix, $\Phi_{(i,j)}(t) = e^{At}_{(i,j)}$ The following theorem is the counterpart to Theorem 1:

Theorem B.1

Assume $A_{(i,j)}$ exists and is continuous at b_o. Then for each $t \in [0,t_f]$, the second partial derivative $\Phi_{(i,j)}(t)$ is given by the integral equation

$$\Phi_{(i,j)}(t) = \int_0^t \Phi(t-s)\{A_{(j)}\Phi_{(i)}(s) + A_{(i)}\Phi_{(j)}(s) + A_{(i,j)}\Phi(s)\}ds \tag{426}$$

or as the unique solution to the matrix differential equation

$$\dot{\Phi}_{(i,j)}(t) = A\Phi_{(i,j)}(t) + A_{(j)}\Phi_{(i)}(t) + A_{(i)}\Phi_{(j)}(t) + A_{(i,j)}\Phi(t) \tag{427}$$

$$\Phi_{(i,j)}(0) = 0 \tag{428}$$

Proof

Let $\Phi_{(i)}(t)$ be the unique solution of

$$\dot{\Phi}_{(i)}(t) = A\Phi_{(i)}(t) + A_{(i)}\Phi(t) \qquad \Phi_{(i)}(0) = 0 \tag{429}$$

and $\Phi_{(i)_{\Delta_j}}(t)$ denote the unique solution of

$$\dot{\Phi}_{(i)_{\Delta_j}}(t) = A_{\Delta_j}\Phi_{(i)_{\Delta_j}}(t) + A_{(i)_{\Delta_j}}\Phi_{\Delta_j}(t) \qquad \Phi_{(i)_{\Delta_j}}(0) = 0 \tag{430}$$

where $A_{\Delta_j} \equiv A(b_o + e_j\Delta b_j)$, etc. Equation (430) may be written as

$$\dot{\Phi}_{(i)_{\Delta_j}}(t) = A\Phi_{(i)_{\Delta_j}}(t) + A_{(i)}\Phi(t) + (A_{\Delta_j} - A)\Phi_{(i)_{\Delta_j}}(t)$$

$$+ A_{(i)}(\Phi_{\Delta_j}(t) - \Phi(t)) + (A_{(i)_{\Delta_j}} - A_{(i)})\Phi_{\Delta_j}(t) \tag{431}$$

which has unique solution

$$\Phi_{(i)_{\Delta_j}}(t) = \int_0^t \Phi(t-s)\{A_{(i)}\Phi(s) + (A_{\Delta_j} - A)\Phi_{(i)_{\Delta_j}}(s) + A_{(i)}(\Phi_{\Delta_j}(s) - \Phi(s)) + (A_{(i)_{\Delta_j}} - A_{(i)})\Phi_{\Delta_j}(s)\}ds \quad (432)$$

Using the above expression and the definition of the second partial derivative

$$\Phi_{(i,j)}(t) \equiv \lim_{|\Delta b_j| \to 0} \frac{\Phi_{(i)_{\Delta_j}}(t) - \Phi_{(i)}(t)}{\Delta b_j} \quad (433)$$

expression (426) of Theorem B.1 readily follows (see proof of Theorem 1). Equation (427) then follows as a straight-forward application of Leibnitz rule.

Q.E.D.

Like the essential role which the augmented matrix A_i plays in the definition of the first order partial derivatives, $e^{At}_{(i)}$, the augmented matrix

$$A_{i,j} \equiv \begin{bmatrix} A & 0 & 0 & 0 \\ A_{(i)} & A & 0 & 0 \\ A_{(j)} & 0 & A & 0 \\ A_{(i,j)} & A_{(j)} & A_{(i)} & A \end{bmatrix} \quad (434)$$

plays an essential role in the definition of the second partial derivatives, $e^{At}_{(i,j)}$:

Theorem B.2

$$e^{A_{i,j}t} = \begin{bmatrix} e^{At} & 0 & 0 & 0 \\ e^{At}_{(i)} & e^{At} & 0 & 0 \\ e^{At}_{(j)} & 0 & e^{At} & 0 \\ e^{At}_{(i,j)} & e^{At}_{(j)} & e^{At}_{(i)} & e^{At} \end{bmatrix} \tag{435}$$

Proof

Using the relation

$$\frac{d}{dt} e^{\tilde{A}_{i,j}t} = \tilde{A}_{i,j} e^{A_{i,j}t} \qquad e^{A_{i,j}0} = I \tag{436}$$

and Theorem 2 and Theorem B.1, it is straight-forward to obtain equation (435) of Theorem B.2. (See proof of Theorem 2.)

Q.E.D.

Like Theorem 2, Theorem B.2 can be used to give fundamental algebraic relations for the second partial derivatives, $e^{At}_{(i,j)}$. However, before stating such representations, the following general result concerning the minimal polynomial of A_{ij} will be useful. The notation used in Section III concerning the minimal polynomial of A will be maintained here.

Lemma B.1

Suppose that the μ-order minimal polynomial of A is

$$\psi(q) = \prod_{k=1}^{\rho} (q - q_k)^{\mu_k} \tag{437}$$

Then the minimal polynomial of A_{ij} is

$$\tilde{\psi}_{ij}(q) = \prod_{k=1}^{\rho} (q - q_k)^{\tilde{\mu}_{ij_k}} \tag{438}$$

and has order $\tilde{\mu}_{ij}$ where

$$\mu \leq \tilde{\mu}_{ij} \leq 3\mu \tag{439}$$

and the multiplicities have similar bounds

$$\mu_k \leq \tilde{\mu}_{ij_k} \leq 3\mu_k \tag{440}$$

Proof

To simplify the notation let $D \equiv \text{adj}\,[qI - A]$; let Δ be the characteristic polynomial of A; and let the unknown matrix components of adj $[qI - \tilde{A}_{ij}]$ be designated

$$\text{adj}[qI - \tilde{A}_{ij}] = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \tag{441}$$

Then it is true that (Ref 21)

$$[qI - A_{ij}]\ \text{adj}\ [qI - \tilde{A}_{ij}] = \Delta^4 I \tag{442}$$

which can be used to form a set of 16 linear equations in terms of the 16 unknown matrix elements of adj$[qI - \tilde{A}_{ij}]$. Then using the relation that

$$[qI - A]D = \Delta I \tag{443}$$

it is straight-forward, but algebraically somewhat tedious, to show that:

$$a_2 = a_3 = a_4 = b_3 = b_4 = c_4 = 0 \tag{444}$$

$$a_1 = b_2 = c_3 = d_4 = \Delta^3 D \tag{445}$$

$$b_1 = d_3 = \Delta^2 DA_{(i)}D \tag{446}$$

$$c_1 = d_2 = \Delta^2 DA_{(j)}D \tag{447}$$

$$d_1 = \Delta^2 DA_{(i,j)}D + \Delta DA_{(j)}DA_{(i)}D + \Delta DA_{(i)}DA_{(j)}D. \tag{448}$$

Now suppose that $g(q)$ is the polynomial such that

$$g(q) = \Delta(q)/\psi(q). \tag{449}$$

Then $g(q)$ is the largest common factor of every term of D (Ref 21), and so the polynomial $\Delta(q)g^3(q)$ is a common factor of every term of $\text{adj}[qI - \tilde{A}_{ij}]$. The minimal polynomial of $\tilde{A}_{ij}$ is given by

$$\tilde{\psi}_{ij}(q) = \Delta^4(q)/g_{ij}(q) \tag{450}$$

where $g_{ij}(q)$ is the largest common factor of every term of $\text{adj}[qI - \tilde{A}_{ij}]$. From the discussion above, $\Delta(q)g^3(q)$ is a factor of $g_{ij}(q)$, and so the results of the Lemma follow.

Q.E.D.

Notice that like Lemma 6, explicit expressions for the minimal polynomial of $\tilde{A}_{ij}$ could be provided by examining common factors of equations (443) - (447). Also, if structural constraints on the minimal polynomial of A are assumed in a neighborhood of $b_o \in R^p$, then stronger conclusions concerning the relationship between the minimal polynomial of $\tilde{A}_{i,j}$ and the eigenvalue sensitivities can be made. However, these stronger results

are notationally cumbersome and have little apparent practical value, and so they are not included here.

Lemma B.1 is now used in combination with Theorem B.2 to give the basic algebraic description of second partial derivatives of e^{At}:

Corollary B.2.1

$$e^{At} = \sum_{k=1}^{3n} A^{k-1} \hat{a}_k(t) \tag{451}$$

$$e^{At}_{(i)} = \sum_{k=1}^{3n} (A^{k-1})_{(i)} \hat{a}_k(t) \tag{452}$$

$$e^{At}_{(i,j)} = \sum_{k=1}^{3n} (A^{k-1})_{(i,j)} \hat{a}_k(t) \tag{453}$$

for all i, j = 1, 2, ...p where the 3n scalar functions $\hat{a}_k(\cdot)$ are uniquely determined from the 3n' equations

$$t^j e^{q_k t} = \frac{d^j}{dq^j} [1 \quad q \ \ldots \ q^{3n-1}]_{q=q_k} \begin{bmatrix} \hat{a}_1(t) \\ \vdots \\ \hat{a}_{3n}(t) \end{bmatrix} \tag{454}$$

$k = 1, 2, \ldots \rho$

$j = 0, 1, \ldots 3n_k - 1$

Proof

It is straight-forward to show that

$$A^k_{ij} = \begin{bmatrix} A^k & 0 & 0 & 0 \\ (A^k)_{(i)} & A^k & 0 & 0 \\ (A^k)_{(j)} & 0 & A^k & 0 \\ (A^k)_{(i,j)} & (A^k)_{(j)} & (A^k)_{(i)} & A^k \end{bmatrix} \tag{455}$$

Then using Theorem B.2 it is easy to see where a polynomial representation of order 4n could be obtained. However, using Lemma B.1, this polynomial representation may be reduced to 3n.

Q.E.D.

This is the central result from which convenient algebraic description of the second order sensitivity operators can be obtained.

Theorem B.3

For the linear time-invariant system S_{LC} it is true that:

$$\text{i)} \quad y_{z.i.}(t) = \sum_{k=1}^{3n} (CA^{k-1}d)\hat{a}_k(t) \tag{456}$$

$$y_{z.s.}(t;u) = \sum_{k=1}^{3n} (CA^{k-1}B) \int_0^t \hat{a}_k(t-s)u(s)ds \tag{457}$$

$$\text{ii)} \quad v^{(i)}_{z.i.}(t) = \sum_{k=1}^{3n} (CA^{k-1}d)_{(i)}\hat{a}_k(t) \tag{458}$$

$$v^{(i)}_{z.s.}(t;u) = \sum_{k=1}^{3n} (CA^{k-1}B)_{(i)} \int_0^t \hat{a}_k(t-s)u(s)ds \tag{459}$$

$$\text{iii)} \quad v^{(i,j)}_{z.i.}(t) = \sum_{k=1}^{3n} (CA^{k-1}d)_{(i,j)}\hat{a}_k(t) \tag{460}$$

$$v^{(i,j)}_{z.s.}(t;u) = \sum_{k=1}^{3n} (CA^{k-1}B)_{(i,j)} \int_0^t \hat{a}_k(t-s)u(s)ds \tag{461}$$

for all i, j = 1, 2, ...p, and where the 3n scalar functions are determined from equation (454).

Proof

The proof is a straight-forward application of Corollary B.2.1 to the definition of the first and second order sensitivity operators.

Q.E.D.

The computational considerations of Theorem B.3 are very similar to those of Theorem 4. Notice that the technique of Appendix A can be used to convert the indicated convolution integrals into a total of at most 3nr quadrature integrals. Also recall from Appendix A that the number of quadrature integrals will automatically be reduced to the minimum number as unnecessary "modes" will be multiplied by zero "component" matrices in the component representation of $e^{At}_{(i,j)}$, and so these unnecessary modes can be eliminated. Finally, we comment that techniques similar to those discussed in Section IV.3 may be applied here to yield a minimal dimension differential equation description of the second order sensitivity system. Such an approach would be an alternative to minimal-order sensitivity models of Bingulac (Ref 34) and Varshney (Ref 40), and would yield the absolute minimum number of differential equations.

BIBLIOGRAPHY

For the convenience of the reader we have divided the Bibliography into major categories of references. Since the dissertation covers a fairly broad range of areas in systems theory, it is hoped that this organization will assist the reader in locating the references for which he is particularly interested. This Bibliography is by no means exhaustive but merely lists the references which are particularly relevant or illustrative to the points made in the body of the dissertation. These categories are the following: A. General References, B. Minimal Order Models for Sensitivity, C. Eigenvalue Sensitivity, D. Insensitivity and Invariance, E. Minimum Sensitivity Control Laws, F. General References on Parameter Identification, G. Identifiability, H. Input Design for Improved Parameter Identification, and I. Computation with Walsh Functions.

A. General References

1. Baras, J. S., R. W. Brockett, and P. A. Fuhrmann. "State Space Models of Infinite Dimensional Systems", IEEE Trans. Auto. Control, AC-19: 693-700, Dec 1974.

2. Bellman, R. Introduction to Matrix Analysis, New York: McGraw Hill, 2nd Edition, 1970.

3. Blum, E. K. Numerical Analysis and Computation Theory and Practices, Reading, Mass: Addison-Wesley, 1972.

4. Bode, H. W. Network Analysis and Feedback Amplifier Design, New York: Van Nostrand, 1945.

5. Bryson, A. E., and Y. C. Ho. Applied Optimal Control Theory, Waltham, Mass: Ginn Company, 1966.

6. Cruz, J. B., Jr., Editor. System Sensitivity Analysis, Stroudsburg, Pa: Dowden Hutchinson Ross, 1973.

7. Cruz, J. B., Jr., and W. R. Perkins. "A New Approach to the Sensitivity Problem in Multivariable Feedback System Design", IEEE Trans. Auto. Control, AC-19: 216-223, 1964.

8. Csaki, F. G. "Some Notes on the Inversion of Confluent Vandermonde Matrices", IEEE Trans. Auto. Control, AC-20: 154-157, Feb 1975.

9. Delfour, M. C. and S. K. Mitter. "Controllability, Observability and Optimal Feedback Control of Affine Hereditary Differential Systems", SIAM J. Control, 10, 2: 298-328, May 1972.

10. Delfour, M. C., and S. K. Mitter. "Controllability and Observability for Infinite Dimensional Systems", SIAM J. Control, 10, 2: 329-333, May 1972.

11. Faddeev, D. K., and V. N. Faddeeva. Computational Methods of Linear Algebra, trans by R. C. Williams, San Francisco: W. H. Freeman Pub., 1963.

12. Fel'dbaum, A. A. Optimal Control Systems, New York: Academic Press, 1965.

13. Goknar, I. C. "Obtaining the Inverse of the Generalized Vandermonde Matrix of the Most General Type", IEEE Trans. Auto. Control, AC-18: 530-532, Oct 1973.

14. Horowitz, I. M. Synthesis of Feedback Systems, New York: Academic Press, 1963.

15. Horowitz, I. M., and U. Shaked. "Superiority of Transfer Function Over State-Variable Methods in Linear Time-Invariant Feedback System Design", IEEE Trans. Auto. Control, AC-20: 84-97, Feb 1975.

16. Kalman, R. E. "Mathematical Description of Linear Dynamical Systems", SIAM J. Control, 1: 152-192, 1963.

17. Kalman, R. E., Y. C. Ho, and K. S. Narendra. "Controllability of Linear Dynamical Systems", Contributions to Differential Equations, 1, 2: 189-213, 1963.

18. Lee, R. C. K. Optimal Estimation, Identification, and Control, M. I. T. Press, 1964.

19. Liusternik, L. A., and W. I. Sobolev. Elements of Functional Analysis, New York: Frederick Ungar Pub. Co., 1961.

20. Luenberger, D. G. Optimization by Vector Space Methods, New York: John Wiley, 1969.

21. Nering, E. D. Linear Algebra and Matrix Theory, New York: John Wiley, 1970.

22. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. The Mathematical Theory of Optimal Processes, (Trans by K. N. Trirogoff, ed. by L. W. Neustadt), New York: John Wiley, 1962.

23. Porter, W. A. "Minimizing System Sensitivity Through Feedback", J. Franklin Institute, 286, 3: 225-229.

24. Rosen, J. "The Gradient Projection Method for Nonlinear Programming. Part I. Linear Constraints", J. Soc. Indus. Appl. Math., 8: 181-217, 1960.

25. Royden, H. L. Real Analysis, New York: The MacMillan Co., 2nd Edition, 1968.

26. Silverman, L. M., and H. E. Meadows. "Controllability and Time-Variable Unilateral Networks", IEEE Trans. Circuit Theory, CT-12: 308-312, Sep 1965.

27. Taylor, A. E. Introduction to Functional Analysis, New York: John Wiley, 1958.

28. Tomovic, R. Sensitivity Analysis of Dynamic Systems, New York: McGraw Hill, 1963.

29. Tomovic, R., and M. Vukobratovic. General Sensitivity Theory, New York: American Elsevier, 1972.

30 Tse, E., and Y. Bar-Shalom. "An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach", IEEE Trans. Auto. Control, AC-18: 109-116, Apr 1973.

31. Van Trees, H. L. Detection, Estimation, and Modulation Theory, Part I, New York: John Wiley, 1968.

32. Weiss, H. "Use of Semi-Groups for Derivation of Controllability and Observability in Time-Invariant Linear Systems", Int. J. Control, 18, 3: 475-479, 1973.

33. Zadeh, L. A., and C. A. Desoer. Linear System Theory, New York: McGraw Hill, 1963.

B. Minimal Order Models for Parameter Sensitivity

34. Bingulac, S. P. "Simultaneous Generation of the Second Order Sensitivity Functions", IEEE Trans. Auto. Control, AC-11: 563-566, Jul 1966.

35. Denery, D. G. "Simplification in the Computation of the Sensitivity Functions for Constant Coefficient Linear Systems", IEEE. Trans. Auto. Control, AC-16, 3. Aug 1971.

36. Guardabassi, G., A. Locatelli, and S. Rinaldi. "On the Optimality of the Wilkie-Perkins Low-Order Sensitivity Model", IEEE Trans. Auto. Control, AC-15: 382-384, Jun 1970.

37. Kokotovic, P. V. "Methods of Sensitivity Points in the Investigation of Linear Control Systems", Automatika i' Telemekhanika, 25, 12: 1670-1676, Dec 1964.

38. Neumann, C. P., and A. K. Sood. "Sensitivity Functions for Multi-Input Linear Time Invariant Systems", Int. J. Control, 13, 6: 1137-1150, 1971.

39. Neumann, C. P., and A. K. Sood. "Sensitivity Functions for Multi-Input Linear Time-Invariant Systems - II. Minimal Order Models", Int. J. Control, 15, 3: 451-463, 1972.

40. Varshney, R. K. "Reduced-Order Sensitivity Models of Linear Systems with Applications", Ph. D. Dissertation, University of Illinois, Urbana, 1972.

41. Wilkie, D. F. and W. R. Perkins. "Essential Parameters in Sensitivity Analysis", Automatica, 5: 191-197, 1969.

42. Wilkie, D. F. and W. R. Perkins. "Generation of Sensitivity Functions for Linear Systems Using Low-Order Models", IEEE Trans. Auto. Control, AC-14: 123-130, 1969.

C. Eigenvalue Sensitivity

43. Crossley, T. R. and B. Porter. "Eigenvalue and Eigenvector Sensitivities in Linear Systems Theory", Int. J. Control, 10, 2: 163-170, 1969.

44. Crossley, T. R. and B. Porter. "High-Order Eigenproblem Sensitivity Methods: Theory and Application to the Design of Linear Dynamical Systems", Int. J. Control, 10, 3: 315-329, 1969.

45. Mantey, P. E. "Eigenvalue Sensitivity and State-Variable Selection", IEEE Trans. Auto. Control, AC-13: 263-269, Jun 1968.

46. Morgan, B. S., Jr. "Sensitivity Analysis and Synthesis of Multivariable Systems", IEEE Trans. Auto. Control, AC-11: 506-512, 1966.

47. Singer, R. A. "Selecting State Variables to Minimize Eigenvalue Sensitivity of Multivariable Systems", Automatica, 5: 85-93, 1969.

D. Sensitivity Controllability and Uncontrollability

48. Bertele, U., and G. Guardabassi. "Can the Terminal Condition of Time-Invariant Linear Control Systems be Made Immune Against Small Parameter Variations?", IEEE Trans. Auto. Control, AC-16: 460-462, Oct 1971.

49. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Structural Uncontrollability of Sensitivity Systems", Automatica, 5: 297-301, May 1969.

50. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Further Results On Controllability and Sensitivity", IEEE Trans. Auto. Control, AC-15: 697-698, Dec 1970.

51. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Structural Properties of Sensitivity Systems", J. Franklin Institute, 294, 4: 241-248, Oct 1972.

52. Guardabassi, G., A. Locatelli, and S. Rinaldi. "The Role of Controllability in Some Sensitivity Problems", Int. J. Control, 19, 1: 57-64, 1974.

53. Sarvaragi, Y., J. Inoue, and K. Asai. "Synthesis of Open-Loop Optimal Control with Zero Sensitivity Terminal Constraints", Automatica, 5: 389-394, May 1969.

E. Insensitivity and Invariance

54. Aggarwal, J. K., H. T. Banks, and N. H. McClamroch. "Invariance in Linear Systems", J. Math. Anal. and Appl., 29: 498-506, 1970.

55. Bonivento, C. "Structural Insensitivity Versus Identifiability", IEEE Trans. Auto. Control, AC-18: 190-192, Apr 1973.

56. Bonivento, C., R. Guidorzi, and G. Marro. "Parameter Insensitivity and Controlled Invariance", Third IFAC Symposium on Sensitivity Adaptivity and Optimality, Ischia, Italy, Jun 18-21, 1973.

57. Cruz, J. B. and W. R. Perkins. "Conditions for Signal and Parameter Invariance in Dynamical Systems", IEEE Trans. Auto. Control, AC-11: 614-615, Jul 1966.

58. Fabian, E. and W. M. Wonham. "Decoupling and Data Sensitivity", IEEE Trans. Auto. Control, AC-20: 338-344, Jun 1975.

59. Fronza, G. and A., Locatelli. "Insensitivity by Linear Feedback", J. Franklin Institute, 296, 4: 237-247, Oct 1973.

60. Karlin, A., A. Locatelli, and C. Zanardini. "Trajectory Insensitivity via Feedback", Automatica, 10: 517-524, 1974.

61. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Insensitivity and Parameteric Invariance in Linear Systems", Int. Rep LCA 71-9, Instituto di Elletrotecnica ed Elettronica, Politecnico di Milano, Apr 1971.

62. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Parameteric Invariance of Linear Systems", Richerche di Automatica, 3, 1, Apr 1972.

63. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Zero-Sensitivity in Linear Systems", Int. J. Control, 15, 4: 785-791, 1972.

64. Guardabassi, G., A. Locatelli, and S. Rinaldi. "Signal Insensitivity, Controllability, and Observability in Linear Systems", IEEE Trans. Auto. Control, AC-16: 277-278, Jun 1970.

65. Karlin, A., A. Locatelli, and C. Zanardini. "Trajectory Insensitivity via Feedback", Automatica, 10: 517-524, 1974.

66. Khrustalev, M. M. "Necessary and Sufficient Conditions for Weak Invariance", Automatika i Telemekhanika, 4: 17-22, Apr 1968.

67. Rozoner, L. I. "A Variational Approach to the Problem of Invariance of Automatic Control Systems", Automatika I Telemekhanika, 24, 6: 744-756, Jun 1963.

F. Minimum Sensitivity Control Laws

68. Guardabassi, G., A. Locatelli, and S. Rinaldi. "On the Optimization of Continuous Linear Systems with Sensitivity Constraints", Preprints Second IFAC Symposium on System Sensitivity and Adaptivity, Dubrovnik, Yugslavia, Aug 1968.

69. Holtzman, J. M., and S. Horing. "The Sensitivity of Terminal Conditions of Optimal Control Systems to Parameter Variations", IEEE Trans. Auto. Control, AC-10: 420-426, Oct 1965.

70. Kahne, S. "Low-Sensitivity Design of Optimal Linear Control Systems", IEEE Trans. Aerospace and Elect. Systems, AES-4: 374-379, May 1968.

71. Kwakernaak, H. "Optimal Low-Sensitivity Linear Feedback Systems", Automatica, 5: 279-285, 1969.

72. Lamont, G., and S. Kahne. "Comparison of Sensitivity Improvement Techniques for Linear Optimal Control Systems", IEEE. Trans. Aerospace and Elect. Systems, AES-5: 142-151, Mar 1969.

73. Luh, J. "Minimum Trajectory Sensitivity of a Large Launch Booster Control System", IEEE Trans. Aerospace and Elect. Systems, AES-5: 213-220, Mar 1969.

74. Pederson, K. C., and L. R. Nardizzi. "Adaptive Control Using Optimally Sensitive Design Techniques", Int. J. Control, 17, 2: 435-447, 1973.

75. Price, C., and J. Deyst. "A Method for Obtaining Desired Sensitivity Characteristics with Optimal Controls", Proceedings Joint Auto. Control Conf., University of Michigan, 1968.

76. Rohrer, R., and M. Sobral, Jr. "Sensitivity Considerations in Optimal System Design", IEEE Trans. Auto. Control, AC-10: 43-48, Jan 1965.

77. Stavroulakis, P., and P. Sarachik. "Low Sensitivity Feedback Gains for Deterministic and Stochastic Control Systems", Int. J. Control, 19, 1: 15-31, 1974.

78. Zadicario, J., and R. Swan. "The Optimal Control of Linear Systems with Unknown Parameters", IEEE Trans. Auto. Control, AC-11: 423-426, Jul 1966.

G. General References on Parameter Identification

79. Allison, J. S. "On the Comparison of Two Methods of Off-Line Parameter Identification", J. Math. Anal. and Appl., 18: 229-237, 1967.

80. Åstrom, K. J., and P. Eykhoff. "System Identification - A Survey", Automatica, 7: 123-162. 1971.

81. Banks, H. T., and G. M. Groome. "Convergence Theorems for Parameter Estimation by Quasilinearization", J. Math. Anal. and Appl., 42: 91-109, 1969.

82. Cuenod, M., and A. P. Sage. "Comparison of Some Methods Used for Process Identification", Automatica, 4: 235-269, 1968.

83. Eykhoff, P. "Process Parameter and State Estimation", Automatica, 4: 205-233, 1968.

84. Gupta, N. K., and R. K. Mehra. "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations", IEEE Trans. Auto. Control, AC-19: 774-783, Dec 1974.

85. Kitzero, R. A. "System Level IMU Calibration", Proceedings of Symposium on A. F. Appl. of Modern Control Theory, Wright-Patterson AFB, Ohio, Jul 1974.

86. Kumar, K. S. P., and R. Sridhar. "On the Identification of Control Systems by the Quasi-Linearization Method", IEEE Trans. Auto. Control, AC-9: 151-154, Apr 1964.

87. Saridis, G. N. "Stochastic Identification Methods for Identification and Control - A Survey", IEEE Trans. Auto. Control, AC-19: 798-809, Dec 1974.

H. Identifiability

88. Aoki, M., and P. C. Yue. "On Certain Convergence Questions in System Identification", SIAM J. Control, 8, 2: 239-256.

89. Bellman, R., and K. J. Åstrom. "On Structural Identifiability", Mathematical Biosciences, 7: 313-328, 1970.

90. Denham, M. J. "Canonical Forms for the Identification of Multi-Variable Linear Systems", IEEE Trans. Auto. Control, AC-19: 646-656, Dec 1974.

91. Glover, K. "Structural Aspects of System Identification", Ph.D. Dissertation, M.I.T., Aug 1973.

92. Glover, K. "Some Geometrical Properties of Linear Systems with Implications in Identification", Submitted to IFAC Congress, Boston, Aug 1975.

93. Glover, K., and J. C. Willems. "Parameterizations of Linear Dynamical Systems: Canonical Forms and Identifiability", IEEE Trans. Auto. Control, AC-19: 640-645, Dec 1974.

94. Hoberock, L. L. and G. W. Stewart. "Input Requirements and Parametric Errors for System Identification Under Periodic Excitation", Trans. of the ASME, 296, Dec 1972.

95. Kim, C. H., and D. P. Lindorff. "Input Frequency Requirements for Identification Through Liapunov Methods", Int. J. Control, 20, 1: 35-48, 1974.

96. Martenson, K. "Least Squares Identifiability of Dynamic Systems", Tech. Rpt. RB73-44, Dept. Elec. Eng., USC, Los Angeles, CA, Aug 1973.

97. Mayne, D. Q. "A Canonical Model for Identification of Multivariable Linear Systems", IEEE Trans. Auto. Control, AC-17: 728-731, Oct 1972.

98. Mehra, R. K. "Identification in Control and Econometrics: Similarities and Differences", Tech. Rpt. No. 647, Div. Eng. and Applied Physics, Harvard Univ., Cambridge, MA, Jul 1973.

99. Payne, R. L., and G. C. Goodwin. "On the Identifiability of Linear Dynamic Systems", Int. J. Control, 20, 5: 865-868, 1974.

100. Popov, V. M. "Invariant Description of Linear Time-Invariant Controllable Systems", SIAM J. Control, 10, 2: 252-264, 1974.

101. Staley, R. M., and P. C. Yue. "On System Parameter Identifiability", Information Sciences, 2: 127-138, 1970.

102. Tse, E. "Information Matrix and Local Identifiability of Parameters", Proceedings Joint Auto. Control Conf., Ohio State Univ., Columbus, OH, 1973.

103. Tse, E., and J. J. Anton. "On the Identifiability of Parameters", IEEE Trans. Auto. Control, AC-17: 637-646, Oct 1972.

I. Input Design for Improved Parameter Identification

104. Aoki, M., and R. M. Staley. "On Input Signal Synthesis in Parameter Identification", Automatica, 6: 431-440, 1970.

105. Barker, H. A., and D. Raeside. "Linear Modeling of Multivariable Systems with Pseudo-Random Binary Input Signals", Automatica, 4: 393-416, 1968.

106. Funahashi, Y., and K. Nakamura. "Identification of Discrete-Time Systems Using a Composite Sinusoidal System", Int. J. Control, 18, 5: 945-962, 1973.

107. Goodwin, G. C., J. C. Murdoch, and R. L. Payne. "Optimal Test Signal Design for Linear S.I.S.O. System Identification", Int. J. Control, 17, 3: 45-55, 1973.

108. Gupta, N. K., and W. E. Hall. "Input Design for Identification of Aircraft Stability and Control Derivatives", NASA-CR-2493, Feb 1975.

109. Kalaba, R. E., and K. Spingarn. "Optimal Inputs and Sensitivities for Parameter Estimation", J. Opt. Theory and Appl., 11, 1: 56-67, 1973.

110. Kalaba, R. E., and K. Spingarn. "Optimal Input System Identification for Homogeneous and Non-Homogeneous Boundary Conditions", J. Opt. Theory and Appl., 16, 5: 487-496, Sep 1975.

111. Levadi, V. S. "Design of Input Signals for Parameter Estimation", IEEE Trans. Auto. Control, AC-11: 205-211, Apr 1966.

112. Mehra, R. K. "Optimal Inputs for Linear System Identification", Preprints, 1972 Joint Auto. Control Conf., Stanford, CA, Aug 1972; also IEEE Trans. Auto. Control, AC-19: 192-200, Jun 1974.

113. Mehra, R. K. "Frequency-Domain Synthesis of Optimal Inputs for Linear System Parameter Estimation", TR645, Division of Engineering and Applied Physics, Harvard Univ., Jul 1973.

114. Mehra, R. K. "Optimal Input Signals for Parameter Estimation in Dynamic Systems - Survey and New Results", IEEE Trans. Auto. Control, AC-19: 753-768, Dec 1974.

115. Mehra, R. K. "Time-Domain Synthesis of Optimal Inputs for System Identification", Preprints IEEE Decision and Control Conf., Phoenix, AZ, Dec 1974.

116. Mehra, R. K. "Frequency-Domain Synthesis of Optimal Inputs for Multiinput-Multioutput (MIMO) Synthesis with Process Noise", Preprints IEEE Decision and Control Conf., Phoenix, AZ, Dec 1974.

117. Nahi, N. E., and G. A. Napjus. "Design of Optimal Probing Signals for Vector Parameter Estimation", Preprints IEEE Decision and Control Conf., Miami, FL, Dec 1971.

118. Nahi, N. E., and D. E. Wallis, Jr. "Optimal Inputs for Parameter Estimation in Dynamic Systems with White Observation Noise", Preprints, Joint Auto. Control Conf., Boulder, CO, Aug 1969.

119. Reid, D. B. "Optimal Inputs for System Identification", Ph.D. Dissertation, Stanford Univ., Stanford, CA, SUDAAR 440, May 1972.

120. Stepner, D. E., and R. K. Mehra. "Maximum Likelihood Identification and Optimal Input Design for Identifying Aircraft Stability and Control Derivatives", NASA Contractor Rpt., NASA CR-2200, Mar 1973.

121. Swanson, G. D., and J. W. Bellville. "Forcing Function Selection for Identification in Human Respiratory Control", Preprints, IEEE Decision and Control Conf., San Diego, CA, Dec 1973.

122. Zarrop, M. B., and G. C. Goodwin. "Comments on Optimal Inputs for System Identification", IEEE Trans. Auto. Control, AC-20: 299-300, Apr 1975.

J. Computation with Walsh Functions

123. Chen, C. F., and C. H. Hsiao. "Walsh Series Analysis in Optimal Control", Int. J. Control, 21, 6: 881-898, Jun 1975.

124. Chen, C. F., and C. H. Hsiao. "Design of Piecewise Constant Gains for Optimal Control via Walsh Functions", IEEE Trans. Auto. Control, AC-20, 5: 596-602, Oct 1975.

125. Corrington, M. S. "Solution of Differential and Integral Equations with Walsh Functions", IEEE Trans. on Circuit Theory, CT-20, 5: 470-476, Sep 1973.

126. Walsh, J. L. "A Closed Set of Orthogonal Functions", American J. Math., 45: 5-24, 1923.

☆ U. S. GOVERNMENT PRINTING OFFICE: 1976 — 757-001/82